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

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

James Isenberg Richard Montgomery

With the scientific input of

Jędrzej Śniatycki Philip B.Yasskin

July 19, 2006

arXiv: physics/9801019
## Contents

1 Introduction

I—Covariant Field Theory

2 Multisymplectic Manifolds 19

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2A</td>
<td>The Jet Bundle</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>22</td>
</tr>
<tr>
<td>2B</td>
<td>The Dual Jet Bundle</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>28</td>
</tr>
</tbody>
</table>

3 Lagrangian Dynamics 31

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3A</td>
<td>The Covariant Legendre Transformation and the</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Cartan Form</td>
<td></td>
</tr>
<tr>
<td>3B</td>
<td>The Euler–Lagrange Equations</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>37</td>
</tr>
</tbody>
</table>

4 Covariant Momentum Maps and Noether’s Theorem 42

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4A</td>
<td>Jet Prolongations</td>
<td>42</td>
</tr>
<tr>
<td>4B</td>
<td>Covariant Canonical Transformations</td>
<td>44</td>
</tr>
<tr>
<td>4C</td>
<td>Covariant Momentum Maps</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>48</td>
</tr>
<tr>
<td>4D</td>
<td>Symmetries and Noether’s Theorem</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>56</td>
</tr>
</tbody>
</table>

Interlude I—On Classical Field Theory 61

II—Canonical Analysis of Field Theories

5 Symplectic Structures Associated with Cauchy Surfaces 67

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5A</td>
<td>Cauchy Surfaces and Spaces of Fields</td>
<td>67</td>
</tr>
<tr>
<td>5B</td>
<td>Canonical Forms on $T^*\gamma_\tau$ and $Z_\tau$</td>
<td>70</td>
</tr>
</tbody>
</table>

---

*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 DMS 96-33161.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5C</td>
<td>Reduction of $\mathcal{L}<em>\tau$ to $T^*Y</em>\tau$</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>74</td>
</tr>
<tr>
<td>6</td>
<td>Initial Value Analysis of Field Theories</td>
<td>76</td>
</tr>
<tr>
<td>6A</td>
<td>Slicings</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>83</td>
</tr>
<tr>
<td>6B</td>
<td>Space + Time Decomposition of the Jet Bundle</td>
<td>85</td>
</tr>
<tr>
<td>6C</td>
<td>The Instantaneous Legendre Transform</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>90</td>
</tr>
<tr>
<td>6D</td>
<td>Hamiltonian Dynamics</td>
<td>96</td>
</tr>
<tr>
<td>6E</td>
<td>Constraint Theory</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>116</td>
</tr>
<tr>
<td>7</td>
<td>The Energy-Momentum Map</td>
<td>124</td>
</tr>
<tr>
<td>7A</td>
<td>Induced Actions on Fields</td>
<td>124</td>
</tr>
<tr>
<td>7B</td>
<td>The Energy-Momentum Map</td>
<td>126</td>
</tr>
<tr>
<td>7C</td>
<td>Induced Momentum Maps on $T^*Y_\tau$</td>
<td>129</td>
</tr>
<tr>
<td>7D</td>
<td>The Hamiltonian and the Energy-Momentum Map</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>Interlude II—The Stress-Energy-Momentum Tensor</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>149</td>
</tr>
<tr>
<td>III</td>
<td>III—Gauge Symmetries and Initial Value Constraints</td>
<td>153</td>
</tr>
<tr>
<td>8</td>
<td>The Gauge Group</td>
<td>153</td>
</tr>
<tr>
<td>8A</td>
<td>Covariance, Localizability, and Gauge Groups</td>
<td>153</td>
</tr>
<tr>
<td>8B</td>
<td>Principal Bundle Construction of the Gauge Group</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>155</td>
</tr>
<tr>
<td>8C</td>
<td>Gauge Transformations</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>160</td>
</tr>
<tr>
<td>9</td>
<td>The Vanishing Theorem and Its Converse</td>
<td>162</td>
</tr>
<tr>
<td>9A</td>
<td>The Vanishing Theorem</td>
<td>162</td>
</tr>
<tr>
<td>9B</td>
<td>The Converse of the Vanishing Theorem</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>Examples</td>
<td>164</td>
</tr>
</tbody>
</table>
V—Palatini Gravity 227

14 Application to Palatini Gravity 227
   14A Covariant Analysis ........................................ 228
   14B Canonical Analysis ........................................ 231
   14C Group-theoretical Analysis ................................. 249
   14D The Adjoint Formalism .................................... 252
   14E The ADM Formalism ....................................... 253

References 254
1 Introduction

This is a five part work in which we study the Lagrangian and Hamiltonian structures 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 this book, 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, 1979b], Isenberg and Nester [1980], Künzle and Nester [1984], Bao, Isenberg, and Yasskin [1985], Batlle et al. [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” \( \psi \) (with conjugate momenta \( \rho \)); 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

\[
\Phi^i(\psi, \rho) = 0
\]

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 $\Phi^i$ 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 $\Phi^i$ to appear in the Hamiltonian. Indeed, in these examples, we find:

4. The Hamiltonian (with respect to a slicing of both spacetime and the relevant bundles of fields over it) has the form

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

(1.2)

depending linearly on the “atlas fields” $\alpha_i$. Here $\Sigma$ is a hypersurface—usually taken to be a Cauchy surface—modeling a typical member of the given spacetime 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 $\mathcal{J}$ be an almost complex structure compatible with both the given inner product and the symplectic form. Let $\lambda$ be a parameter (“time”) labeling the slicing mentioned above. From (1.2), it follows that
5. The evolution equations for the dynamic fields \((\psi, \rho)\) take the “adjoint form”

\[
\frac{d}{d\lambda} \left( \begin{array}{c} \psi \\ \rho \end{array} \right) = J \cdot \sum_i \left[ D\Phi^i(\psi(\lambda), \rho(\lambda)) \right]^* \alpha_i(\lambda).
\]

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 \textit{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 et al. [1980], Arms et al. [1981, 1982], Marsden [1988], Arms [1977, 1981], Arms et al. [1990], Arms et al. [1991], sjamaar and Lerman [1991], and Interlude IV 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 our 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 $\Phi^i$ 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 [1993], Deligne and Freed [1999], 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 et al. [1976], Gotay et al. [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 et al. [1962]) language, the superhamiltonian $\mathcal{H}$ and supermomenta $\mathcal{J}$ 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 $\mathcal{H}$ and $\mathcal{J}$ 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 decomposition shows that the covariant momentum map induces an “energy-momentum” map $\Phi$ on the instantaneous phase space. It is not a genuine momentum map, as there is no group action in this context; nonetheless, $\Phi$ 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, $\Phi = -(\mathcal{H}, \mathcal{J})$, so that the superhamiltonian and supermomenta are the components of the energy-momentum map.

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] and Martínez [2004]. This is, however, a subject for future research.)
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 \textit{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:

1. a covariant analysis of field theories;

2. a space + time decomposition of the covariant formalism followed by an initial value analysis of field theories;

3. a study of the relations between gauge symmetries and initial value constraints; and

4. the 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 part of this book, starting with the covariant analysis. The fifth part in this book 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 III–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 Chapter 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 \textit{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 $\Theta$. The multisymplectic $(n+2)$-form $\Omega = -d\Theta$ is the covariant generalization of the symplectic form in Hamiltonian mechanics. The pair $(Z, \Omega)$ is called multiphase space.

Next, in Chapter 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 $\mathbb{F}\mathcal{L}: J^1Y \to Z$ which is defined, as in mechanics, by means of the fiber derivative of $\mathcal{L}$. Pulling the form $\Omega$ 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], Ragonieri 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, 1991b] 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 et al. [2003]) in relation to taking “horizontal variations” of the action.

In Chapter 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 $G$ (possibly infinite-dimensional) with Lie algebra $\mathfrak{g}$ that acts on $(Z, \Omega)$ in an appropriate way, then one can define a multi- or covariant momentum map $\mathcal{J}: Z \to \mathfrak{g}^* \otimes \Lambda^n Z$ that intertwines the group action with the multisymplectic structure via the equation

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

where $\langle J, \xi \rangle$ denotes $J$ paired with $\xi \in \mathfrak{g}$, $\xi_Z$ is the infinitesimal generator on $Z$. 
of the one-parameter group generated by $\xi$, and $d$ is the exterior differential. 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 $Y_\Sigma$ of sections of $Y$ restricted to $\Sigma$, and the phase space is $T^*Y_\Sigma$ 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 Chapter 5 is that the multisymplectic structure on $Z$ induces the canonical symplectic structure on $T^*Y_\Sigma$. To bridge the gap between the covariant and instantaneous formalisms, we introduce an intermediate space $Z_\Sigma$, the space of sections of $Z$ restricted to $\Sigma$. The multisymplectic $(n + 2)$-form $\Omega$ on $Z$ induces a presymplectic 2-form $\Omega_\Sigma$ on $Z_\Sigma$ by integration over $\Sigma$. Reducing $Z_\Sigma$ by the kernel of $\Omega_\Sigma$ produces a symplectic manifold $Z_\Sigma/\ker\Omega_\Sigma$, which we prove is canonically symplectomorphic to $T^*Y_\Sigma$.

This shows how to space + time decompose covariant multisymplectic phase spaces into instantaneous symplectic phase spaces. Next, in §§6A–6D, we perform a similar decomposition of dynamics using the notion of slicings. An
important result here 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 et al. [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 Chapter 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 $\mathcal{Y}_\Sigma$ or on $T^*\mathcal{Y}_\Sigma$ as it need not stabilize $\Sigma$. Nonetheless, by integrating over $\Sigma$, the covariant momentum map $J$ gives rise to a map $\mathcal{L}_\Sigma \to \mathfrak{g}^*$, which can then be shown to project to a well-defined map $\mathcal{E}_\Sigma : \mathcal{P}_\Sigma \to \mathfrak{g}^*$, where the “primary constraint set” $\mathcal{P}_\Sigma$ is the image of $T^\ast \mathcal{Y}_\Sigma$ in $T^\ast T^\ast \mathcal{Y}_\Sigma$ under the instantaneous Legendre transform. This map $\mathcal{E}_\Sigma$ is the energy-momentum map. It is typically not a true momentum map. However, if one considers only the subgroup $\mathcal{G}_\Sigma$ of $\mathcal{G}$ that stabilizes the Cauchy surface (and that therefore acts in the instantaneous formalism), then $\mathcal{E}_\Sigma$ restricts to a genuine momentum map for the $\mathcal{G}_\Sigma$-action on $T^\ast \mathcal{Y}_\Sigma$ that corresponds in relativity to the supermomenta, without the inclusion of the superhamiltonian. In the general case $\mathcal{E}_\Sigma$ corresponds to both $\mathcal{G}$ and $\mathcal{J}$.

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 estab-
lishing 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 Chapter 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.

For a given field theory, a **gauge group** $\mathcal{G}$ is a subgroup of $\text{Aut}(Y)$, the group of automorphisms of the configuration bundle $Y$. It must satisfy certain properties, viz.:

- $\mathcal{G}$ acts by **symmetries**, for example, the Lagrangian density $L$ is $\mathcal{G}$-equivariant.
- $\mathcal{G}$ 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.
- $\mathcal{G}$ is “sufficiently large” (in a sense to be made precise in Part III).

In §8B 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 $\text{Aut}(Y) \to \text{Diff}(X)$. 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 and that all fields are variational. (Although our formalism is best suited to such theories, much of it—with appropriate modifications—will apply to “background” theories as well as parametrized theories in which some fields are nonvariational. These points are discussed further in the text and the examples; 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 $\Sigma$ and an element $\zeta \in \mathfrak{g}$ with the property that its associated spacetime vector field $\zeta_X$ is transverse to $\Sigma$. (Such a vector field exists since the theory is parametrized.) The Lie algebra element $\zeta$ may be thought of as defining an “evolution direction”; it is analogous to the “lapse” and “shift” in relativity. In §7D we show that the instantaneous Hamiltonian density $H_{\Sigma,\zeta}$ on $\mathcal{P}_\Sigma$ corresponding to the evolution direction $\zeta \in \mathfrak{g}$ is given by

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

(1.4)

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

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

Equation (1.4) forms the basis for item 4; in this regard, the fact that $H_{\Sigma,\zeta}$ is manifestly linear in $\zeta$ presages its linearity in the atlas fields, discussed later in Chapter 13 for general field theories.

Next, in Chapter 9, we prove the Vanishing Theorem (aka the “Second Noether Theorem”), which states that when evaluated on any solution of the Euler–Lagrange equations, the multimomentum map integrated over a hypersurface is zero. This result follows from Noether’s Theorem, $\mathfrak{g}$-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 $\mathcal{E}_\Sigma$ must vanish, as opposed to merely being constant, is a direct consequence of localizability.)

With these results in hand, in Chapter 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 $\mathcal{C}_\Sigma$ of admissible Cauchy data for the evolution equations is exactly the zero level set of the energy-momentum
map associated to the gauge group:

\[ \mathcal{C}_\Sigma = \mathcal{E}_\Sigma^{-1}(0). \]  

(1.5)

The set \(\mathcal{C}_\Sigma\) 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 \(\mathcal{G}\) is nonzero, 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 Chapter 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 $\mathcal{E}_\Sigma$ is only defined on the primary constraint set $\mathcal{P}_\Sigma \subset T^*\mathcal{Y}_\Sigma$. In Chapter 11 we show that (first class) primary constraints can be recovered in much the same way, but the proof is surprisingly more difficult and requires different techniques. For the primaries the relevant object is not $\mathcal{G}$ but rather a certain foliation $\mathcal{J}_\Sigma$ of $\mathcal{Y}_\Sigma$ derived from $\mathcal{G}$; intuitively, $\mathcal{J}_\Sigma$ consists of the “time derivatives” of elements of the subgroup $\mathcal{G}_\Sigma$ which stabilizes $\Sigma$. This foliation has a “momentum map” $\mathcal{J}_\Sigma$ and, provided zero is a regular value, the primary constraints are obtained by equating $\mathcal{J}_\Sigma$ 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 $\mathcal{G}_\Sigma$; 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 $\mathcal{T}^{\mu\nu}$ 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 $\mathcal{X}$, our $\mathcal{T}^{\mu\nu}$ 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 Chapter 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 $D_\Sigma = Y_\Sigma / \dot{G}_\Sigma$. 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 $\dot{G}_\Sigma$, so their canonically conjugate fields are parametrized by $\dot{G}_\Sigma$.

By a slight generalization of the cotangent bundle reduction theorem (Abraham and Marsden [1978], Kummer [1981], and Marsden [1992]), the instantaneous phase space $T^* Y_\Sigma$ can then be cut down to the space $T^* D_\Sigma$ of dynamic fields $\psi$ and their conjugate momenta $\rho$. (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 $P_\Sigma$ in $T^* Y_\Sigma$ to parametric dynamics on $T^* D_\Sigma$. The parameters will be certain combinations of the kinematic fields and elements of $g$ which we call “atlas fields.”

Unfortunately, the energy-momentum map $\mathcal{E}_\Sigma$ on $P_\Sigma$ does not directly project to $T^* D_\Sigma$, as it involves the kinematic fields. But $\mathcal{E}_\Sigma$, thought of as a map $P_\Sigma \times g \to \mathbb{R}$, also depends upon $\xi \in g$. By algebraically lumping the kinematic fields together with elements of the Lie algebra of the gauge group in a “$\dot{G}_\Sigma$-equivariant” fashion, one forms the nondynamic **atlas fields** $\alpha_i$. 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 $\mathcal{E}_\Sigma$ depends only upon $(\psi, \rho)$ and the $\alpha_i$. The energy-momentum map then can be pushed down to a function $\Phi_\Sigma$ on $T^* D_\Sigma$ depending upon the parameters $\alpha_i$. This function $\Phi_\Sigma$ is the “reduced” energy-momentum map. We observe that the Energy-Momentum Theorem remains valid in this context, since the secondary constraints restrict
only \((\psi, \rho)\)—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 \textit{linear} in the atlas fields. This is vital for the \textit{adjoint form} (1.3) of the evolution equations, which now follows readily from (1.2) by writing Hamilton’s equations explicitly on \(T^* \mathcal{D}\).

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 \(\alpha_i\) (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 Chapter 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. We provide a brief overview of these topics in Interlude IV.

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 (see Šniatycki [1974] and §10C); this presents no problem in principle, and is usually straightforward 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, 1991b]; 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 $G$-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 et al. [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 et al. [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-Lopez and Marsden [2003] and Śniatycki [2004].

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 et al. [1998] have shown that the treatment of water waves and their instabilities by Bridges [1997] may be understood in multisymplectic terms and that one can build numerical integrators using these ideas. In Interlude III we survey some of the basic ideas of variational (or multisymplectic) integrators, so the reader may consult that section for further information. One can also do continuum mechanics from the multisymplectic point of view as in Marsden et al. [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 et al. [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, Wiktor Szczyrba, and Garth Warner 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,
§1 Introduction

Annapolis, Berkeley, Boston, Boulder, Brno, Canberra, Champaign-Urbana, Clausthal-Zellerfeld, College Station, Eugene, Hamburg, Hämelschenburg, Honolulu, Houston, Irvine, 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 chapter 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], Garcúa [1974], Kijowski and Szczyrba [1975, 1976], Kijowski and Tulczyjew [1979], Aldaya and de Azcárraga [1980a], Kupershmidt [1980], Kosmann-Schwarzbach [1981], Ragonieri and Ricci [1981], Kastrup [1983], Śniatycki [1970a, 1984], de León and Rodrigues [1985], Günther [1987], Gotay [1991a, 1991b], Sardanashvily [1993], Kanatchikov [1998], Marsden et al. [1998], and Castrillon-Lopez and Marsden [2003]. 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 $C^\infty$. Our conventions follow Misner et al. [1973] to a large extent.

2A The Jet Bundle

Let $X$ be an oriented manifold, which in many examples is spacetime, and let $\pi_{XY} : Y \to X$ 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 section $\phi_1, \phi_2$ 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 $T_xX \to T_yY$. 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 $\gamma : T_xX \to T_yY$ satisfying

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

The map $\gamma$ corresponds to $T_x\phi$ for a local section $\phi$. 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

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

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

$$T_y Y = \text{im} \gamma \oplus V_y Y. \quad (2A.3)$$

**Remark 2A.1.** One can view sections of $J^1 Y$ over $Y$ as *Ehresmann connections* on $Y$. (See Hermann [1975] or Marsden et al. [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 $J^1 Y \to Y$ is the bundle whose equivariant sections are connections in the usual sense of principal bundles.\(^2\)

In Chapter 3 we shall study Lagrangians on $J^1 Y$ 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 $J^1 Y$ 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 de Azcárraga [1980b], Shadwick [1982], Ferraris and Francaviglia [1983], Krupka [1987], Horáček and Kolář [1983], de León and Rodrigues [1985], Gotay [1991a, 1991b], and references therein.

We let $\dim X = n + 1$ and the fiber dimension of $Y$ be $N$. Coordinates on $X$ are denoted $x^\mu$, $\mu = 0, 1, 2, \ldots, n$, and fiber coordinates on $Y$ are denoted by $y^A$, $A = 1, \ldots, N$. These induce coordinates $v^A_\mu$ on the fibers of $J^1 Y$. If $\phi : X \to Y$ is a section of $\pi_{XY}$, 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 $J^1 Y$ regarded as a bundle over $X$. This section is denoted $j^1 \phi$ and is called the *first jet prolongation* of $\phi$. In coordinates, $j^1 \phi$ is given by

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

where $\partial_\nu = \partial/\partial x^\nu$. A section of the bundle $J^1 Y \to X$ which is the first jet of a section of $Y \to X$ 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.

---

\(^2\) In this case $J^1 Y$, as a principal bundle over the connection bundle $CY = J^1 Y/G$, itself has a canonical connection. For the bundle $Q \times \mathbb{R} \to Q$ 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 $\theta$. The curvature of this connection is the canonical symplectic form $d\theta$, up to a sign depending on conventions.
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 \( \pi_{XY} \) the projection onto the first factor. The first jet bundle \( J^1Y \) is the bundle whose holonomic sections are tangents of sections \( \phi : X \to Y \) (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 \( \Lambda^1X \) of one-form potentials \( A \). Coordinates for \( J^1Y \) are \((x^\mu, A_\nu, v^\nu_\mu)\), where the derivative index \( \mu \) 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 \( J^1\tilde{Y} \), respectively, where \( S_2^{3,1}(X) \) is the bundle of symmetric covariant symmetric two-tensors of Lorentz signature \((-,-,+,+\)).

The coordinates on \( J^1\tilde{Y} \) are now \((x^\mu, A_\nu, v^\nu_\mu; g_\rho_\nu, w^\nu_\rho_\mu)\), where we use a semicolon to separate variational fields from parametric ones.\(^3\)

\(^3\) Henceforth we use a tilde to distinguish objects in parametrized electromagnetism from
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 $J^1 Y$ 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 $x^\mu$ for $\mu = 0, 1$. Also, let $(M, g)$ be a $(d + 1)$-dimensional spacetime manifold $M$ with coordinates $\phi^A$ for $A = 0, 1, \ldots, d$ and a fixed metric $g$ of Lorentz signature $(-, +, ..., +)$. A **bosonic string** is a map $\phi : X \to M$. 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 $S^1_2(X)$ denote the bundle over $X$ of symmetric covariant rank two tensors of Lorentz signature $(-, +)$. Then a section of $S^1_2(X)$ 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]), $\phi$ and $h$ are taken as independent variables and the Lagrangian is taken as the harmonic map Lagrangian. In the second (due to Nambu [1970]), $\phi$ is taken as the independent variable, while $h$ is taken to be the induced metric

$$h_{\sigma\rho}(x) = (\phi^A g)_{\sigma\rho}(x) = g_{AB}(\phi(x)) \phi^A,_{\sigma}(x) \phi^B,_{\rho}(x).$$

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 $\phi$ 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 book since it treats the metric $h$ variationally. This model has also been treated by Batlle et al. [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] and Green et al. [1987]. One can analyze the Nambu model in a similar way; see, for example, Scherk [1975].

With $\phi$ and $h$ chosen as the independent variables, the covariant configuration bundle is $Y = (X \times M) \times_X S^{1,1}_2(X)$ and

$$J^1Y = J^1(X \times M) \times_X J^1(S^{1,1}_2(X)).$$

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

---

**Remark 2A.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 $S^{3,1}_2(X)$. 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 et al. [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 et al. [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.
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 $y \in Y$ is the set of affine maps from $J^1_y Y$ to $\Lambda^{n+1}_x X$, where $\Lambda^{n+1}_x X$ denotes the bundle of $(n+1)$-forms on $X$. A smooth section of $J^1Y^*$ is therefore an affine bundle map of $J^1Y^*$ to $\Lambda^{n+1}_x X$ covering $\pi_{XY}$.

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

$$v^A_\mu \mapsto (p + p_A^\mu v^A_\mu) d^{n+1}x$$

(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 $\Lambda_y$ and with projection $\pi_{Y\Lambda} : \Lambda \to Y$. Let $Z \subset \Lambda$ be the subbundle whose fiber is given by

$$Z_y = \{ z \in \Lambda_y | i_v i_w z = 0 \text{ for all } v, w \in V_y Y \} ,$$

(2B.2)

where $i_v$ denotes left interior multiplication by $v$.

Elements of $Z$ can be be written uniquely as

$$z = p d^{n+1}x + p_A^\mu dy^A \wedge d^n x_\mu$$

(2B.3)

where

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

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

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

$$\Phi : Z \to J^1Y^*.$$
Intrinsically, $\Phi$ is defined by
\[ \langle \Phi(z), \gamma \rangle = \gamma^* z \in \Lambda_x^{n+1}X \tag{2B.4} \]
where $z \in Z_y$, $\gamma \in J^1_y Y$, $x = \pi_{XY}(y)$, and $\langle \cdot, \cdot \rangle$ denotes the dual pairing. To see this, note that if $\gamma$ has fiber coordinates $v^\mu$, then
\[ \gamma^* dx^\mu = dx^\mu \quad \text{and} \quad \gamma^* dy^A = v^\mu dx^\mu \tag{2B.5} \]
and so
\[ \gamma^* (pd^{n+1}x + pA^\mu dy^A \wedge d^n x_\mu) = (p + pA^\mu v^\mu) d^{n+1}x, \tag{2B.6} \]
where we have used
\[ dx^\nu \wedge d^n x_\mu = \delta^\nu_\mu d^{n+1}x. \]

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

We summarize:

\textbf{Proposition 2B.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 \textit{canonical} $(n+1)$-\textit{form} $\Theta_\Lambda$ on $\Lambda$ by
\[ \Theta_\Lambda(z)(u_1, \ldots, u_{n+1}) = z(T\pi_{Y\Lambda} \cdot u_1, \ldots, T\pi_{Y\Lambda} \cdot u_{n+1}) \]
\[ = (\pi^*_Y z)(u_1, \ldots, u_{n+1}), \tag{2B.7} \]
where $z \in \Lambda$ and $u_1, \ldots, u_{n+1} \in T_z \Lambda$. Define the \textit{canonical} $(n + 2)$-\textit{form} $\Omega_\Lambda$ on $\Lambda$ by
\[ \Omega_\Lambda = -d\Theta_\Lambda. \tag{2B.8} \]
Note that if $n = 0$ (i.e., $X$ is one-dimensional), then $\Lambda = T^*Y$ and $\Theta_\Lambda$ is the standard canonical one-form. If $i_{AZ} : Z \to \Lambda$ denotes the inclusion, the \textit{canonical} $(n + 1)$-\textit{form} $\Theta$ on $Z$ is defined by
\[ \Theta = i^*_{AZ} \Theta_\Lambda \tag{2B.9} \]
and the \textit{canonical} $(n + 2)$-\textit{form} $\Omega$ on $Z$ is defined by
\[ \Omega = -d\Theta = i^*_{AZ} \Omega_\Lambda. \tag{2B.10} \]
The pair \((Z, \Omega)\) is called **multiphase space** or **covariant phase space**. It is an example of a **multisymplectic manifold**; see Remark 2B.4 below.

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

\[
\Theta = p_A^\mu dy^A \wedge d^n x_\mu + p d^{n+1} x,
\]

and so

\[
\Omega = dy^A \wedge dp_A^\mu \wedge d^n x_\mu - dp \wedge d^{n+1} x.
\]

**Remark 2B.2.** There are other possible choices of multiphase space (cf. the paper by Echeverria-Enríques et al. [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 2B.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.

**Remark 2B.3.** A notable feature of the covariant formalism is the appearance of several momenta \(p_A^\mu\), spatial in addition to temporal, for each field component \(y^A\). The Legendre transformation of Chapter 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.

**Remark 2B.4.** 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 \(\Omega\) which is nondegenerate in the sense that \(i_V \Omega \neq 0\) 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 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...
We next recall that there is a characterization of the canonical one-form $\theta$ on $T^*Q$ by $\beta^*\theta = \beta$ for any one-form $\beta$ on $Q$ (Abraham and Marsden [1978] Proposition 3.2.11). Here is an analogue of that fact for the canonical $(n+1)$-form.

**Proposition 2B.5.** If $\sigma$ is a section of $\pi_{XZ}$ and $\phi = \pi_{YZ} \circ \sigma$, then

$$\sigma^*\Theta = \phi^*\sigma,$$

(2B.13)

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

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

$$(\sigma^*\Theta)(x)(v_1, \ldots, v_{n+1}) = \Theta(\sigma(x))(T\sigma \cdot v_1, \ldots, T\sigma \cdot v_{n+1})$$

$$= \sigma(x)(T\pi_{YZ} \cdot T\sigma \cdot v_1, \ldots, T\pi_{YZ} \cdot T\sigma \cdot v_{n+1})$$

$$= \sigma(x)(T\phi \cdot v_1, \ldots, T\phi \cdot v_{n+1})$$

$$= (\phi^* \sigma)(x)(v_1, \ldots, v_{n+1}).$$

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 $(t, p, q^1, \ldots, q^N, p_1, \ldots, p_N)$,

$$\Theta = p_A dq^A + p \, dt$$

(2B.14)

and

$$\Omega = dq^A \wedge dp_A + dt \wedge dp.$$  

(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 gothic notation in anticipation. From (2B.11) and (2B.12), we get

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

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 $\tilde{Z}$ with coordinates $(x^\mu, A_\nu, p, \mathfrak{F}^{\nu\mu}; g_{\sigma\rho}, \rho^{\sigma\rho})$. The expressions for the canonical forms are modified accordingly:

$$\tilde{\Theta} = (\mathfrak{F}^{\nu\mu} dA_\nu + \rho^{\sigma\rho} dg_{\sigma\rho}) \wedge d^3x_\mu + pd^4x$$

and

$$\tilde{\Omega} = (dA_\nu \wedge d\mathfrak{F}^{\nu\mu} + d\rho^{\sigma\rho} \wedge dg_{\sigma\rho}) \wedge d^3x_\mu - dp \wedge d^4x.$$

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

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

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 $(x^\mu, \phi^A, h_{\sigma\rho}, p, p_{A\mu}, \rho^{\sigma\rho})$, and the canonical forms are given by

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

and

$$\Omega = (d\phi^A \wedge dp_{A\mu} + dh_{\sigma\rho} \wedge d\rho^{\sigma\rho}) \wedge d^1x_\mu - dp \wedge d^2x. \quad (2B.21)$$
We conclude this section with a few remarks on Poisson brackets. Poisson brackets are not essential for the main points in this book, but we believe that they are important for linking up our results with those of Marsden et al. [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-Lopez and Marsden [2003], and Forger et al. [2003a, 2003b].

Recall that on a symplectic manifold \((P, \Omega)\), the Hamiltonian vector field \(X_f\) of a function \(f \in \mathcal{F}(P)\) is defined by
\[
\Omega(X_f, V) = V \lhd df
\]
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 \(\pi_{XY}\)-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
\[
\frac{f(z)}{= \pi_{YZ}^*(V \lhd z)} \tag{2B.22}
\]
for some \(\pi_{XY}\)-projectable vector field \(V\); we call such an \(f\) a \textit{momentum observable}. (As we shall see in \S 4C, special covariant momentum maps can be regarded as momentum observables.) The \textit{Hamiltonian vector field} \(X_f\) of such a momentum observable \(f\) is defined by
\[
\frac{df}{=} = X_f \lhd \Omega
\]
where \(\Omega\) is the canonical \((n + 2)\)-form on \(Z\). Since \(\Omega\) is nondegenerate, this uniquely defines \(X_f\). The \textit{Poisson bracket} of two such \(n\)-forms is the \(n\)-form defined by
\[
\{f, h\} = X_h \lhd (X_f \lhd \Omega). \tag{2B.23}
\]
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 Chapter 5 shows that when equation (2B.23) is integrated over a hypersurface, it produces a canonical cotangent bundle bracket.

3 Lagrangian Dynamics

This chapter 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], Szczryba [1976a], Sternberg [1977], Kupershmidt [1980], Aldaya and de Azcárraga [1980a], Griffiths [1983], Anderson [1992], and Deligne and Freed [1999]. 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, 1991c]. 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 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 Chapter 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 and the Cartan Form

Let
\[ \mathcal{L} : J^1Y \to \Lambda^{n+1}X, \]
the Lagrangian density, be a given smooth bundle map over \( X \). In coordinates, we write
\[ \mathcal{L} = L(x^\mu, y^A, v^A_\mu) \, d^{n+1}x. \tag{3A.1} \]
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
\[ \mathcal{F} \mathcal{L} : J^1Y \to J^1Y^* \cong Z \]
over \( Y \) which has the coordinate expressions
\[ p^A_\mu = \frac{\partial L}{\partial v^A_\mu}, \quad p = L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu, \tag{3A.2} \]
for the multimomenta \( p^A_\mu \) and the covariant Hamiltonian \( p \). An intrinsic definition follows.

If \( \gamma \in J^1_yY \) then \( \mathcal{F} \mathcal{L}(\gamma) \) is the affine approximation to \( \mathcal{L} \bigr|_{J^1_yY} \) at \( \gamma \). Thus \( \mathcal{F} \mathcal{L}(\gamma) \) is the affine map from \( J^1_yY \) to \( \Lambda_x^{n+1}X \) (where \( y \in Y_x \)) given by
\[ \langle \mathcal{F} \mathcal{L}(\gamma), \gamma' \rangle = \mathcal{L}(\gamma) + \frac{d}{d\varepsilon} \mathcal{L}(\gamma + \varepsilon(\gamma' - \gamma)) \biggr|_{\varepsilon=0}, \tag{3A.3} \]
where \( \gamma' \in J^1_yY \). 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 \( w^A_\mu \) 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 \( \mathcal{F} \mathcal{L} \). (The coordinatization of \( J^1Y^* \) was defined by (2B.1).)
§3B The Euler–Lagrange Equations

The **Cartan form** is the \((n+1)\)-form \(\Theta_L\) on \(J^1Y\) defined by

\[
\Theta_L = F^*L^*\Theta,
\]

(3A.4)

where \(\Theta\) is the canonical \((n+1)\)-form on \(Z\). We also define the \((n+2)\)-form \(\Omega_L\) by

\[
\Omega_L = -d\Theta_L = F^*L^*\Omega,
\]

(3A.5)

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

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

\[
\Theta_L = \frac{\partial L}{\partial v^A_\mu} dy^A \wedge d^n x_\mu + \left( L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \right) d^{n+1}x
\]

(3A.6)

and

\[
\Omega_L = dy^A \wedge d\left( \frac{\partial L}{\partial v^A_\mu} \right) \wedge d^n x_\mu - d\left( L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \right) \wedge d^{n+1}x.
\]

(3A.7)

One application of the Cartan form is to re-express the Lagrangian for holonomic fields as

\[
\mathcal{L}(j^1\phi) = (j^1\phi)^*\Theta_L.
\]

(3A.8)

To see this, we use (3A.6) to compute

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

The first and last terms cancel and so we obtain (3A.8).

It is possible to obtain \(\Theta_L\) 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 et al. [1998].

### 3B The Euler–Lagrange Equations

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

\[
\frac{\partial L}{\partial y^A} (j^1\phi) - \frac{\partial}{\partial x^\mu} \left( \frac{\partial L}{\partial v^A_\mu} (j^1\phi) \right) = 0
\]

(3B.1)

for a (local) section \(\phi\) of \(Y\). The left hand side of (3B.1) is often denoted \(\delta L/\delta \phi^A\) and is called the **Euler–Lagrange** or **variational derivative** of \(L\).
Theorem 3B.1. The following assertions regarding a section $\phi$ of the bundle $\pi_{XY} : Y \to X$ are equivalent:

(i) $\phi$ is a stationary point of the action integral $\int_X L(j^1\phi)$;

(ii) the Euler–Lagrange equations (3B.1) hold in coordinates;

(iii) for any vector field $W$ on $J^1Y$,

$$(j^1\phi)^*(W \cup J^1\Omega_L) = 0.$$ (3B.2)

Remark 3B.2. The theorem still holds if the section $\phi$ is only a local section: $\phi : U \to Y_U$ 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 when

$U$ is a neighborhood of a Cauchy surface.

Remark 3B.3. The meaning of “stationary point” is in the formal sense of the calculus of variations. Namely, a variation of $\phi$ is provided by a curve $\phi_\lambda = \eta_\lambda \circ \phi$, where $\eta_\lambda$ is the flow of a vertical vector field $V$ on $Y$ which is compactly supported in $X$. One says that $\phi$ is a stationary point if

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

for all variations $\phi_\lambda$ of $\phi$.

Remark 3B.4. This theorem appears more or less as we present it in Dedecker [1953], Trautman [1967], Śniatycki [1970a, 1984], Ouzilou [1972], Goldschmidt and Sternberg [1973], and Guillemin and Sternberg [1977], to name just a few references.

Remark 3B.5. If in (3B.2) one requires that $s^*(i_W \Omega_L) = 0$ for any section $s$ of $J^1Y \to X$, 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 $(J^1Y, \Omega_L)$ 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 [1991] and Saunders [1992].
Remark 3B.6. 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 et al. [1998].

Remark 3B.7. 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 et al. [1998] and Yavari et al. [2006] for details, and Lew et al. [2003] for further applications of the idea of splitting horizontal and vertical variations.

In the course of the proof of Theorem 3B.1, we will examine three cases for the vector field \( W \) in (3B.2). The following lemma shows that two of the cases are trivial.

Lemma 3B.8. If \( \phi \) is a section of \( \pi_{XY} \) and if either \( W \) is tangent to the image of \( j^1\phi \) or \( W \) is \( \pi_{Y,j^1Y} \)-vertical, then

\[
(j^1\phi)^* (W \cup \Omega_L) = 0.
\]

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

\[
(j^1\phi)^* (W \cup \Omega_L) = (j^1\phi)^* ((T(j^1\phi) \cdot w) \cup \Omega_L)
\]

\[
= w \cup (j^1\phi)^* \Omega_L,
\]

which vanishes since \((j^1\phi)^* \Omega_L\) is an \((n + 2)\)-form on the \((n + 1)\)-manifold \( X \).

Secondly, if \( W \) is \( \pi_{Y,j^1Y} \)-vertical, then \( W \) has the coordinate form

\[
W = (0, 0, W^A_{\mu}).
\]

A calculation using (3A.7) shows that

\[
W \cup \Omega_L = -W^A_{\mu} \frac{\partial^2 L}{\partial v^A_{\mu} \partial v^B_{\nu}} (dy^A \wedge d^n x_{\mu} - v^A_{\mu} d^{n+1} x),
\]

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

Proof of Theorem 3B.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 \( \phi \) (as in Remark 3B.3) corresponding to a \( \pi_{XY} \)-vertical vector field \( V \) on \( Y \) with compact support in \( X \). Using (3A.8), we compute

\[
\frac{d}{d\lambda} \left[ \int_X L(j^1\phi_\lambda) \right]_{\lambda=0} = \frac{d}{d\lambda} \left[ \int_X (j^1\phi_\lambda)^* \Theta_L \right]_{\lambda=0}
\]

\[
= \frac{d}{d\lambda} \left[ \int_X (j^1\phi)^*(j^1\eta_\lambda)^* \Theta_L \right]_{\lambda=0}
\]

\[
= \int_X (j^1\phi)^* \mathcal{L}_{j^1V} \Theta_L,
\]

(3B.3)

where

\[
j^1V = \frac{d}{d\lambda} j^1\eta_\lambda \bigg|_{\lambda=0}
\]

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

\[
j^1V = \begin{pmatrix} 0, V^A, \frac{\partial V^A}{\partial x^\mu} + \frac{\partial V^A}{\partial y^B} v^B_\mu \end{pmatrix}
\]

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 4A.1). Thus, from (3B.3) and

\[
\mathcal{L}_{j^1V} \Theta_L = -j^1V \mathbin{\bigtriangledown} \Omega_L + d(j^1V \mathbin{\bigtriangledown} \Theta_L),
\]

we get

\[
\frac{d}{d\lambda} \left[ \int_X L(j^1\phi_\lambda) \right]_{\lambda=0} = -\int_X (j^1\phi)^*(j^1V \mathbin{\bigtriangledown} \Omega_L) + \int_X d(j^1\phi)^*(j^1V \mathbin{\bigtriangledown} \Theta_L)
\]

\[
= -\int_X (j^1\phi)^*(j^1V \mathbin{\bigtriangledown} \Omega_L)
\]

(3B.4)

by Stokes’ theorem and the fact that \( V \), and hence \( j^1V \) is compactly supported in \( X \). From (3B.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
\(\pi_{X, J^1Y}\)-vertical vector field. Similarly, any \(\pi_{X, J^1Y}\)-vertical vector field can be decomposed into a jet extension of some \(\pi_{XY}\)-vertical vector field \(V\) on \(Y\) and a \(\pi_{Y, J^1Y}\)-vertical vector field.

Thus, if \((i)\) holds, \((3B.4)\) and Lemma \(3B.8\) show that

\[
\int_X (j^1\phi)^*(W \cup \Omega_L) = 0
\]

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 \(j^1\phi\),

\[
(j^1\phi)^* (j^1V \cup \Omega_L) = V^A \left[ \frac{\partial}{\partial x^\mu} \left( \frac{\partial L}{\partial \dot{q}^A} (j^1\phi) \right) - \frac{\partial L}{\partial q^A} (j^1\phi) \right] d^{n+1}x
\]

for all \(\pi_{XY}\)-vertical vector fields \(V\) on \(Y\). Thus \((iii)\) implies \((ii)\). On the other hand, this equation combined with the above remarks on decompositions of vector fields and Lemma \(3B.8\) shows that \((ii)\) implies \((iii)\).

**Remark 3B.9.** This proof establishes the key relation

\[
(j^1\phi)^* (W \cup \Omega_L) = -W^A \frac{\delta L}{\delta \dot{q}^A} d^{n+1}x \quad (3B.5)
\]

for any vector field \(W\) on \(J^1Y\).

---

**Examples**

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

\[
\Theta_L = \frac{\partial L}{\partial \dot{q}^A} dq^A - E dt \quad (3B.6)
\]
and the Euler–Lagrange equations take the standard form

\[
\frac{d}{dt} \frac{\partial L}{\partial v^A} = \frac{\partial L}{\partial q^A}.
\] (3B.7)

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

\[
\mathcal{L} = -m\|v\| dt,
\] (3B.8)

where \(v = (v^0, v^1, v^2, v^3)\) is the velocity of the particle (with respect to parameter time \(t\)) and \(\|v\| = \left(-g_{AB}v^A v^B\right)^{1/2}\), \(g\) being the spacetime metric.\(^5\) The Cartan form (3B.6) is thus

\[
\Theta_L = m g_{AB} v^B \|v\| dq^A,
\] (3B.9)

whence

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

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

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

where \(t\) is used as the curve parameter.

**b Electromagnetism.** For electromagnetism on a fixed background spacetime \(X\) with metric \(g\), the Lagrangian density is\(^6\)

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

\(^5\) Sometimes in this context one encounters the Lagrangian density

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

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

\(^6\) For convenience, in this example we use (3B.12), with the coefficient \(-1/4\), instead of the Misner et al. [1973] Lagrangian, with coefficient \(-1/16\pi\). One may pass from one convention to the other by scaling the electromagnetic fields.
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

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

The Legendre transformation gives the relations

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

The Cartan form $\Theta_L$ is

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

and the 5-form $\Omega_L = -d\Theta_L$ is

$$\Omega_L = \sqrt{-g} (g^{\nu\sigma} g^\mu\rho - g^{\nu\rho} g^{\mu\sigma}) dA_\nu \wedge dv_\rho \wedge d^3x_\mu \quad (3B.16)$$

$$- \sqrt{-g} F^{\nu\mu} dv_{\nu\mu} \wedge d^4x. \quad (3B.17)$$

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

$$(A^{\nu,\mu} - A^{\mu,\nu})_{,\nu} = 0. \quad (3B.18)$$

Here the semicolon denotes a covariant derivative with respect to the background metric, so $(3B.18)$ is a covariant divergence.

When $g$ is treated parametrically, the Legendre transformation gives in addition

$$\rho^{\rho\mu} = 0. \quad (3B.5)$$

(These relations express the fact that the metric does not derivatively couple to the electromagnetic fields in the Lagrangian density.) None of the formulas above change, except $(3B.17)$ for $\Omega_L$ which will pick up terms involving the differentials of the metric. However, since $g$ is not variational, it cannot have Euler–Lagrange equations. In order for this to be the case, we must now interpret the latter to mean:

$$j^1(A; g)^*(W \downarrow \hat{\Omega}_L) = 0$$

for all vector fields $W$ along $J^1(\Lambda^1X)$. (That is, $W$ can have no components tangent to the fibers of $J^1(\Sigma_2^{3,1}(X)) \to X$.) This follows from $(3B.5)$, which

---

7 This is an exceptional circumstance, and is due to the antisymmetry of the field tensor. A Klein–Gordon vector field, for example, would couple derivatively to a metric.
here takes the form
\[ j^1(A; g) (W \cup \Omega_L) = - \left( W_\nu \frac{\delta L}{\delta A_\nu} + W_{\sigma\rho} \frac{\delta L}{\delta g_{\sigma\rho}} \right) d^4x. \]

Thus no conditions are placed on the variational derivatives \( \delta L/\delta g_{\sigma\rho} \) of the metric provided \( W_{\sigma\rho} = 0 \).

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
\[ \mathcal{L} = \frac{1}{2} \epsilon^{\mu\nu\sigma} F_{\mu\nu} A_\sigma d^3x, \quad (3B.19) \]
where the curvature \( F \) of \( A \) is defined by (3B.13). Intrinsically, \( \mathcal{L} = F \wedge A \). Applying the Legendre transformation, we get
\[ p^{\nu\mu} = \epsilon^{\mu\nu\sigma} A_\sigma \quad \text{and} \quad p = 0. \quad (3B.20) \]
The Cartan form is thus
\[ \Theta_{\mathcal{L}} = \epsilon^{\mu\nu\sigma} A_\sigma dA_\nu \wedge d^2x_{\mu}, \quad (3B.21) \]
whence
\[ \Omega_{\mathcal{L}} = - \epsilon^{\mu\nu\sigma} dA_\sigma \wedge dA_\nu \wedge d^2x_{\mu}. \quad (3B.22) \]
The Euler–Lagrange equations are \( \epsilon^{\mu\nu\sigma} A_{[\nu,\mu]} = 0 \), which are equivalent to the connection \( A \) being flat:
\[ (j^1A)^* F = dA = 0. \quad (3B.23) \]
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 \( \phi \) 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
\[ \mathcal{L} = - \frac{1}{2} \sqrt{-h} h^{\sigma\rho} g_{AB} v^A_\sigma v^B_\rho d^2x. \quad (3B.24) \]
The Legendre transformation is

\[ p_A^\mu = -\sqrt{-h} h^{\mu\nu} g_{AB} v^B_\nu \]  \hspace{1cm} (3B.25)

\[ \rho^{p\mu} = 0 \]  \hspace{1cm} (3B.26)

\[ p = \frac{1}{2} \sqrt{-h} h^{\mu\nu} g_{AB} v^A_\mu v^B_\nu, \]  \hspace{1cm} (3B.27)

so that the Cartan form is

\[ \Theta_L = \sqrt{-h} \left( -h^{\mu\nu} g_{AB} v^B_\nu d\phi^A \wedge d^1 x_\mu + \frac{1}{2} h^{\mu\nu} g_{AB} v^A_\mu v^B_\nu d^2 x \right) \]  \hspace{1cm} (3B.28)

and

\[ \Omega_L = d\phi^A \wedge d \left( -\sqrt{-h} h^{\mu\nu} g_{AB} v^B_\nu \right) \wedge d^1 x_\mu \]  \hspace{1cm} (3B.29)

\[- \left( \frac{1}{2} \sqrt{-h} h^{\mu\nu} g_{AB} v^A_\mu v^B_\nu \right) \wedge d^2 x. \]  \hspace{1cm} (3B.30)

The Euler–Lagrange equations \( \delta L/\delta \phi^A = 0 \) and \( \delta L/\delta h^{\alpha\beta} = 0 \) reduce to

\[ (h^{\mu\nu} g_{AB}(\phi) \phi^B_\nu)_{,\mu} = 0 \]  \hspace{1cm} (3B.31)

\[ \left( \frac{1}{2} h^{\mu\nu} g_{AB}(\phi) \phi^A_\mu \phi^B_\nu \right) h_{\alpha\beta} = g_{CD}(\phi) \phi^C_\alpha \phi^D_\beta \]  \hspace{1cm} (3B.32)

respectively. Equation (3B.31) says that \( \phi \) satisfies the harmonic map equation. Equation (3B.32) does two things:

(i) it says \( h \) is conformally related to the induced metric \( \phi^* g \) 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 (3B.32) 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 (3B.32) says that in any coordinate system \( h \) is proportional to \( \phi^* g \). There is no need for a diffeomorphism.
We summarize the content of (3B.32): There are three equations. They say that \( h \) and \( \phi^*g \) are conformal and determine the proportionality constant. But only two of them are independent; indeed, taking the trace of (3B.32), we obtain

\[
\frac{\delta L}{\delta h_{\alpha\beta}} h_{\alpha\beta} = 0.
\]  

(3B.33)

\section{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 \( G \) acting on \( Z \) by covariant canonical transformations. Third, we prove that the lifted action of a group \( G \) 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 \( F_L \). Then we prove Noether’s theorem, here formulated in terms of covariant momentum maps.

\subsection{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 \( \eta_Y : Y \to Y \) be a \( \pi_{XY} \)-bundle automorphism covering a diffeomorphism \( \eta_X : X \to X \). If \( \gamma : T_xX \to T_yY \) is an element of \( J^1Y \), let \( \eta_{J^1Y}(\gamma) : T_{\eta_X(x)}X \to T_{\eta_Y(y)}Y \) be defined by

\[
\eta_{J^1Y}(\gamma) = T\eta_Y \circ \gamma \circ T\eta_X^{-1}.
\]  

(4A.1)
The $\pi_{Y,X}$-bundle automorphism $j^1\eta_Y := \eta_{j^1X}$ so constructed is called the \textit{first jet extension} or \textit{prolongation} of $\eta_Y$ to $J^1Y$. In coordinates, if $\gamma = (x^\mu, y^A, v^A_\mu)$, then

$$\eta_{j^1Y}(\gamma) = \left(\eta^{\mu}_X(x), \eta^{A}_X(x, y), \left[\partial_\nu \eta^{A}_Y + \left(\partial_B \eta^{A}_Y \right)v^B_\nu \right] \partial_\mu (\eta_X^{-1})^\nu \right), \quad (4A.2)$$

where $\partial_B = \partial/\partial y^B$. Occasionally we are interested in the case when $\eta_Y$ is vertical; that is, $\eta_X$ 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 $\eta_\lambda$, so that

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

then its \textit{first jet extension} or \textit{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 $\phi$ is a section of $\pi_{XY}$ and $\eta_Y$ is a $\pi_{XY}$-bundle automorphism covering $\eta_X$, then $\eta_Y \cdot \phi := \eta_Y \circ \phi \circ \eta_X^{-1}$ is another section and

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


\textbf{Remark 4A.1.} If $\phi_\lambda = \eta_\lambda \circ \phi$ is a variation of $\phi$ (as in Remark 3B.3), then $\eta_X$ 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 $\eta_\lambda$ is vertical and (4A.4) reduces to

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

\textbf{Remark 4A.2.} In the above mentioned references, an additional term, namely, $v^B_\mu v^A_\lambda \partial V^\lambda / \partial y^B$ 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 $V^\lambda$ is independent of $y^B$.  ♦
**Remark 4A.3.** A word of caution: One can also view $V$ as a section of the bundle $TY \to Y$ and take its first jet in the sense of equation (2A.4). Then one gets the section of $J^1(TY) \to Y$ given in coordinates by

$$\begin{pmatrix} x^\mu, y^A, V^\mu, \frac{\partial V^\mu}{\partial x^\nu}, \frac{\partial V^\mu}{\partial y^B}, \frac{\partial V^A}{\partial x^\nu}, \frac{\partial V^A}{\partial y^B} \end{pmatrix}.$$

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

**Remark 4A.4.** From the definitions one finds $j^1([V,W]) = [j^1V,j^1W]$.

### 4B Covariant Canonical Transformations

A **covariant canonical transformation** is a $\pi_{YZ}$-bundle map $\eta_Z : Z \to Z$ covering a diffeomorphism $\eta_X : X \to X$ such that

$$\eta_Z^* \Omega = \Omega. \quad (4B.1)$$

We say $\eta_Z$ is a **special covariant canonical transformation** if

$$\eta_Z^* \Theta = \Theta. \quad (4B.2)$$

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

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

that is,

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

where $z \in Z_y$ and $v_1, \ldots, v_{n+1} \in T_{\eta_Y(y)}Y$. Since $\eta_Y$ is a $\pi_{XY}$-bundle map, $T\eta_Y$ maps vertical vectors to vertical vectors, so $\eta_Z$ 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

$$\eta_Z(z) = (\eta_Y^{-1})^* z = \left( p + \partial_\nu (\eta_Y^{-1} A p_A^\mu \partial_\mu \eta_X^\nu \right) J^{-1} d^{n+1}x$$

$$+ \left( \partial_A (\eta_Y^{-1} B p_B^\nu \partial_\nu \eta_X^\mu \right) J^{-1} dy^A \wedge d^n x_\mu, \quad (4B.4)$$
where \( J \) is the Jacobian determinant of \( \eta_X \) and where the quantities in (4B.4) are evaluated at the appropriate points.

**Proposition 4B.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 \( \pi_{XY} \)-bundle automorphism \( \eta_Y : Y \to Y \) covering a diffeomorphism \( \eta_X : X \to X \), which proceeds via \( J^1Y^* \) rather than directly to \( Z \). We define the lift of \( \eta_Y \) to \( J^1Y^* \) to be the affine dual of the prolongation \( \eta_{J^1Y} \); that is, if \( z : J^1_yY \to \Lambda^{n+1}_xX \) is an element of \( J^1_yY^* \), let

\[
\eta_{J^1Y^*}(z) : J^1_{\eta_Y(y)}Y \to \Lambda^{n+1}_{\eta_X(z)}X
\]

be the affine map defined by

\[
\langle \eta_{J^1Y^*}(z), \gamma \rangle = (\eta^{-1}_X)^* \langle z, \eta^{-1}_{J^1Y} (\gamma) \rangle.
\]

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

\[
\langle \eta_{J^1Y^*}(z), \gamma \rangle = (\eta^{-1}_X)^* \left[ \left( p + \sum_A p^A \partial_\mu \eta^\nu_Y \frac{\partial (\eta^{-1}_X)^A}{\partial y^\nu_B} \right) d^{n+1}x \right]
\]

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 4B.2.** The canonical isomorphism \( \Phi : Z \to J^1Y^* \) is equivariant with respect to the lifts \( \eta_Z \) and \( \eta_{J^1Y^*} \); that is,

\[
\Phi \circ \eta_Z = \eta_{J^1Y^*} \circ \Phi
\]

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 \( \gamma \in J^1Y \), we have

\[
\langle \eta_{J^1Y^*}(\Phi(z)), \gamma \rangle = (\eta^{-1}_X)^* \langle \Phi(z), \eta^{-1}_{J^1Y}(\gamma) \rangle
\]

\[
= (\eta^{-1}_X)^* \langle \Phi(z), T\eta_Y^{-1} \circ \gamma \circ T\eta_X \rangle
\]

\[
= (\eta^{-1}_X)^* \left[ (T\eta_Y^{-1} \circ \gamma \circ T\eta_X)^* z \right]
\]

\[
= (\eta^{-1}_X)^* \eta_X^* (\eta^{-1}_Y)^* z = \gamma^* (\eta_Z(z)) = \langle \Phi (\eta_Z(z)), \gamma \rangle
\]
and so (4B.7) holds. Thus, the identification of $Z$ and $J^1Y^\ast$ respects the action of a lifted covariant canonical transformation.

Let $V$ be a vector field on $Y$ whose flow $\eta_\lambda$ maps fibers of $\pi_{XY}$ 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

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

in coordinates, where $V = (V^\mu, V^A)$. By either direct calculation or by Proposition 4B.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 et al. [1983], Guillemin and Sternberg [1984], Marsden [1992], and Marsden and Ratiu [1999]).

Let $G$ 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 $\pi_{XZ}$- (or $\pi_{XY}$-) bundle automorphisms. For $\eta \in G$, let $\eta_X, \eta_Y$, and $\eta_Z$ denote the corresponding transformations of $X, Y$, and $Z$, and for $\xi \in \mathfrak{g}$, let $\xi_X, \xi_Y$, and $\xi_Z$ denote the corresponding infinitesimal generators. If $G$ acts on $Z$ by covariant canonical transformations, then the Lie derivative of $\Omega$ along $\xi_Z$ is zero:

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

and if $G$ acts by special covariant canonical transformations, then

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

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

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

covering the identity on $Z$ such that

$$dJ(\xi) = i_{\xi_Z} \Omega, \quad (4C.3)$$

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

$$J \left( \text{Ad}_q^{-1} \xi \right) = \eta^\ast_Z \left[ J(\xi) \right]. \quad (4C.4)$$
If $G$ acts by *special* covariant canonical transformations, there is an explicit formula for a *special covariant momentum map*. We claim that

$$J(\xi) = i_{\xi_Z} \Theta$$  \hspace{1cm} (4C.5)

is a covariant momentum map. Indeed,

$$dJ(\xi) = di_{\xi_Z} \Theta = (L_{\xi_Z} - i_{\xi_Z} d) \Theta = i_{\xi_Z} \Omega.$$

**Proposition 4C.1.** *Special covariant momentum maps are $\text{Ad}^*$-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 4C.2.** If the action of $G$ on $Z$ is the lift of an action of $G$ on $Y$, then the special covariant momentum map is given by

$$J(\xi)(z) = \pi_{YZ}^* i_{\xi_Y} z.$$  \hspace{1cm} (4C.6)

This momentum map is $\text{Ad}^*$-equivariant.

This follows from the recognition that $\xi_Y = T \pi_{YZ} \cdot \xi_Z$, the definitions of $J(\xi)$ and $\Theta$, and Proposition 4C.1.

In coordinates, if $\xi_Y$ has components $(\xi^\mu, \xi^A)$, then equation (4C.6) reads

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

where

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

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 4C.3.** For lifted actions,

$$\{J(\xi), J(\zeta)\} = d(i_{\xi_Z} i_{\xi_Z} \Theta) + J([\xi, \zeta]).$$  \hspace{1cm} (4C.8)
§4 Covariant Momentum Maps and Noether’s Theorem

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

$$J((\xi, \zeta)) = L_{\xi} J(\xi) = d\iota_{\zeta} J(\xi) + i_{\xi} dJ(\xi) = d\iota_{\zeta} J(\xi) + i_{\xi} i_{\xi} \Omega$$

by (4C.3). This, (4C.5), and (2B.23) give (4C.8). \[\blacksquare\]

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 $G = \text{Diff}(\mathbb{R})$, the diffeomorphism group of $X = \mathbb{R}$. Then $G$ 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

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

is defined by

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

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 $G = \mathcal{F}(X)$, the additive group of smooth functions on spacetime. Let $G$ act on $Y = \Lambda^1 X$ as follows: For $f \in \mathcal{F}(X)$ and $A \in \Lambda^1_x X$,

$$f \cdot A = A + df(x) \in \Lambda^1_x X. \quad (4C.10)$$
For $\chi \in \mathfrak{g} \cong \mathcal{F}(X)$, the infinitesimal generator is given by

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

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

$$\langle J(x^\mu, A_\nu, p, \tilde{\mathcal{F}}^{\mu\nu}), \chi \rangle = \tilde{\mathcal{F}}^{\mu\nu} \chi_\nu d^3 x_\mu. \tag{4C.12}$$

If we treat the metric $g$ on $X$ parametrically, we may enlarge the group $\mathfrak{g}$ to the semi-direct product $\tilde{\mathfrak{g}} = \text{Diff}(X) \ltimes \mathcal{F}(X)$, with $\text{Diff}(X)$ acting on $\mathcal{F}(X)$ by push-forward; that is, $\eta \cdot f = f \circ \eta^{-1}$. Then $\tilde{\mathfrak{g}}$ acts on $\tilde{Y} = Y \times_X S^3_1(X)$ by

$$(\eta, f) \cdot (x, A, g) = (\eta(x), (\eta^{-1})^*(A + df(x)); (\eta^{-1})^*g). \tag{4C.13}$$

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

$$(\xi S^1_2(g))_\sigma^\nu \frac{\partial}{\partial g_\sigma^\nu} + (\xi Y(A))_\nu \frac{\partial}{\partial A_\nu} + (\chi Y(A))_\nu \frac{\partial}{\partial A_\nu} + \xi^\mu \frac{\partial}{\partial x_\mu},$$

where

$$(\xi S^1_2(g))_\sigma^\nu = -(g_\sigma^\mu \xi^\mu, \rho + g_\rho^\nu \xi^\mu, \sigma), \tag{4C.14}$$

$$(\xi Y(A))_\nu = -A_\nu \xi^\mu, \tag{4C.15}$$

and $\chi Y(A)$ is given by (4C.11). From (4C.7) the covariant momentum map $\tilde{J} : \tilde{Z} \to \tilde{\mathfrak{g}}^* \otimes \Lambda^3 \tilde{Z}$ for this action is

$$\langle \tilde{J}(x^\mu, A_\nu, p, \tilde{\mathcal{F}}^{\mu\nu}, g_\sigma^\rho, \rho_\sigma^{\rho\mu}), (\xi, \chi) \rangle =$$

$$\langle -\tilde{\mathcal{F}}^{\mu\nu} A_\tau \xi^{\tau, \nu} + \tilde{\mathcal{F}}^{\mu\nu} \chi_\nu - 2 \rho^{\rho\mu} g_{\sigma\nu} \xi_\rho, \sigma + p \xi^\mu \rangle d^3 x_\mu$$

$$- \xi^\nu (\tilde{\mathcal{F}}^{\tau\mu} dA_\tau + \rho^{\rho\mu} dg_\rho) \wedge d^2 x_{\mu\nu}. \tag{4C.16}$$

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

$$(\eta, f) \cdot A = (\eta^{-1})^*(A + df(x)). \tag{4C.17}$$

---

8 In fact, we could also include conformal rescalings of the metric. This turns out not to add anything essential; moreover, conformal invariance is lost when a gravitational coupling is introduced. We therefore refer the reader to Example d instead for a treatment of conformal rescalings. See also Remark 8C.5.
The corresponding infinitesimal generator is
\[
(\xi, \chi)_Y(A) = (\xi_Y(A))_\nu \frac{\partial}{\partial A_\nu} + (\chi_Y(A))_\nu \frac{\partial}{\partial x^\nu} + \xi^\mu \frac{\partial}{\partial x^\mu},
\]
where \((\xi_Y(A))_\nu\) and \(\chi_Y(A)\) are given by (4C.15) and (4C.11), respectively. Then we compute
\[
\langle J(x^\mu, A_\nu, p, p^{\mu\nu}), (\xi, \chi) \rangle = \left(-p^{\mu\nu} A_\tau \xi^\tau_{\nu} + p^{\nu\mu} \chi^\nu_{\mu} + p \xi^\mu\right) d^2 x_\mu - p^{\tau\mu} \xi^\tau_{\mu} dA_\tau \wedge d^1 x_{\mu\nu}. \tag{4C.19}
\]

**d Bosonic Strings.** Corresponding to \(Y = (X \times M) \times_X S^{1,1}_2(X)\), we take the covariance group for bosonic strings to be the semi-direct product
\[
\mathcal{G} = \text{Diff}(X) \ltimes \mathcal{F}(X, \mathbb{R}^+),
\]
where we view \(\mathcal{F}(X, \mathbb{R}^+)\) as the group of conformal rescalings of \(S^{1,1}_2(X)\) and the relabeling symmetry \(\eta \in \text{Diff}(X)\) acts on \(\Lambda \in \mathcal{F}(X, \mathbb{R}^+)\) by push-forward:
\[
\eta \cdot \Lambda = \Lambda \circ \eta^{-1}. \tag{4C.20}
\]
Then \((\eta, \Lambda) \in \mathcal{G}\) maps \((\phi, h) \in Y_x\) to
\[
(\eta, \Lambda) \cdot (\phi, h) = (\phi, (\eta^{-1})^* (\Lambda^2(x)h)) \tag{4C.21}
\]
in \(Y_{\eta(x)}\). Hence, for \((\xi, \lambda) \in \mathfrak{g} \cong \mathcal{X}(X) \times \mathcal{F}(X)\), the corresponding infinitesimal generator \((\xi, \lambda)_Y\) at \((\phi, h)\) is
\[
(\lambda_Y(h))_{\sigma^\rho} \frac{\partial}{\partial h_{\sigma^\rho}} + (\xi_Y(h))_{\sigma^\rho} \frac{\partial}{\partial h_{\sigma^\rho}} + \xi^\mu \frac{\partial}{\partial x^\mu}, \tag{4C.22}
\]
where
\[
(\lambda_Y(h))_{\sigma^\rho} = 2\lambda h_{\sigma^\rho} \tag{4C.23}
\]
and
\[
(\xi_Y(h))_{\sigma^\rho} = -\left(h_{\sigma^\mu} \xi^\mu_{\,\rho} + h_{\rho^\mu} \xi^\mu_{\,\sigma}\right). \tag{4C.24}
\]
Note that there is no component in the direction \(\partial/\partial \phi^A\) since \(\phi\) is a scalar field on \(X\). From (4C.7), we get the covariant momentum map
\[
\langle J(x^\mu, \phi^A, h_{\sigma^\rho}, p, p_{A}^{\mu}, p^{\rho\nu}), (\xi, \lambda) \rangle = \left[p^{\rho\nu} (2\lambda h_{\sigma^\rho} - h_{\sigma^\mu} \xi^\mu_{\,\rho} - h_{\rho^\mu} \xi^\mu_{\,\sigma}) + p \xi^\mu\right] d^1 x_\mu - (p_{A}^{\mu} \xi^A + p^{\rho\nu} \xi^\rho dA_\tau h_{\sigma^\rho}) \epsilon_{\mu\nu} \tag{4C.25}
\]
where we have written \(\epsilon_{\mu\nu} = d^2 x_{\mu\nu}\).
Remark 4C.4. As we shall see in Part III the covariant momentum maps (4C.12), (4C.19), and (4C.25) are closely related to the divergence, spatial flatness, and superhamiltonian and supermomentum constraints of electromagnetism, Chern–Simons theory, and string theory, respectively.

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–4C, 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 Chapters 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 §9B.

Suppose that a group $G$ 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 $L$ is equivariant with respect to $G$ if for all $\eta \in G$ and $\gamma \in J^1_xY$,

$$L(\eta J^1Y(\gamma)) = (\eta^{-1}_x)^*L(\gamma),$$

where $(\eta^{-1}_x)^*L(\gamma)$ means that the $(n+1)$-form $L(\gamma)$ at $x \in X$ is pushed forward to an $(n+1)$-form at $\eta x$. 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):

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

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

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 $\pi_{XY}$-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 4c below. This assumption has a number of important consequences, some of which are stated in the following proposition.

**Proposition 4D.1.** Let $\mathcal{L}$ be equivariant with respect to the lifted action of $\mathcal{G}$ on $J^1Y$. Then

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

\[
\begin{array}{ccc}
J^1Y & \xrightarrow{F\mathcal{L}} & Z \\
\eta_{J^1Y} & & \eta_Z \\
J^1Y & \xleftarrow{\mathcal{L}} & Z \\
\end{array}
\]

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

\[
\xi_{J^1Y} \hookleftarrow \Omega_L = dJ^\mathcal{L}(\xi).
\]

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

\[
J^\mathcal{L}(\xi) = \xi_{J^1Y} \hookleftarrow \Theta_{\mathcal{L}}.
\]

**Proof.** (i) It is convenient to identify $Z$ with $J^1Y^*$ and use Proposition 4B.2. From (4B.5) and (3A.3) we have

\[
\langle \eta_{J^1Y}, (F\mathcal{L}(\gamma)) \gamma' \rangle = (\eta_X^{-1})^* \langle F\mathcal{L}(\gamma), \eta_{J^1Y}^{-1}(\gamma') \rangle
\]

\[
= (\eta_X^{-1})^* \left[ \mathcal{L}(\gamma) + \frac{d}{d\epsilon} \mathcal{L} (\gamma + \epsilon [\eta_{J^1Y}^{-1}(\gamma') - \gamma]) \bigg|_{\epsilon=0} \right].
\]
On the other hand, using (3A.3) we have

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

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

\[
\xi_Z \circ F_L = TF_L \circ \xi_{J^1Y}.
\]  

Pulling (4C.3) back from Z to J^1Y along the Legendre transformation and using (4D.5) yields (4D.3). Similarly, pulling (4D.5) back by FL yields (4D.4).

Because of this theorem, we call J^L the \textit{covariant momentum map} in the \textit{Lagrangian representation}. From (4D.4) and (3A.6) or, equivalently, from (4C.7) and (3A.2), we have the explicit formula

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

If \( \phi \) is a solution of the Euler–Lagrange equations, then from Theorem 3B.1,

\[
(j^1\phi)^* (W \downarrow \Omega_L) = 0
\]

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 \textit{Noether conservation law}:

\textbf{Theorem 4D.2 (Divergence Form of Noether’s Theorem).} \textit{If the covariance assumption A1 is satisfied, then for each} \( \xi \in \mathfrak{g} \),

\[
d \left[ (j^1\phi)^* J^L(\xi) \right] = 0 \tag{4D.7}
\]

\textit{for any section} \( \phi \) \textit{of} \( \pi_{XY} \) \textit{satisfying the Euler–Lagrange equations}.

The quantity \((j^1\phi)^* J^L(\xi)\) is called the \textit{Noether current}, and Theorem 4D.2 is often referred to as the \textit{first Noether theorem}.\footnote{The “second Noether theorem,” which we call the “Vanishing Theorem,” will be discussed in Chapter 9.} 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
\[
(j^1\phi)^*J^L(\xi) = \left( \frac{\partial L}{\partial v^A} (j^1\phi)(\xi^A \circ \phi) + L(j^1\phi)\xi^\mu - \frac{\partial L}{\partial v^A_{\nu}} (j^1\phi) \phi^A_{,\nu} \xi^\mu \right) d^n x_{\mu} - \frac{\partial L}{\partial v^A_{\mu}} (j^1\phi) \phi^A_{,\lambda} \xi^\nu d^\lambda \land d^{n-1} x_{\mu\nu}
\]
which, upon using the identity
\[
dx_{\lambda} \land d^{n-1} x_{\mu\nu} = \delta_{\nu}^{\lambda} d^n x_{\mu} - \delta_{\mu}^{\lambda} d^n x_{\nu},
\]
simplifies to
\[
(j^1\phi)^*J^L(\xi) = \left[ -\frac{\partial L}{\partial v^A} (j^1\phi)(\xi^A \circ \phi - \phi^A_{,\nu} \xi^\nu) + L(j^1\phi)\xi^\mu \right] d^n x_{\mu}.
\]
(4D.8)

Defining the **Lie derivative** of \( \phi \) along \( \xi \) by
\[
\mathcal{L}_\xi \phi = T\phi \circ \xi + \xi Y \circ \phi; \ \text{i.e.,} \ \ (\mathcal{L}_\xi \phi)^A = \phi^A_{,\nu} \xi^\nu - \xi^A \circ \phi, \quad (4D.9)
\]
(4D.8) may be rewritten as
\[
(j^1\phi)^*J^L(\xi) = \left[ -\frac{\partial L}{\partial v^A} (j^1\phi)(\mathcal{L}_\xi \phi)^A + L(j^1\phi)\xi^\mu \right] d^n x_{\mu} \quad (4D.10)
\]
where \( \phi \) is any section of \( \pi_{XY} \), not necessarily satisfying the Euler–Lagrange
equations.

Next, taking the exterior derivative of the Noether current (4D.8) and using
the identity
\[
d(V^\mu d^n x_{\mu}) = \partial_{\rho} V^\mu d^\rho \land d^n x_{\mu} = \partial_{\rho} V^\mu d^{n+1} x
\]
(which justifies the name “divergence” for the resulting quantity), we obtain
after some rearrangement and cancellation,
\[
\begin{align*}
\mathbf{d} \left[ (j^1 \phi)^* J^\xi(\xi) \right] &= \partial_\mu \left[ \frac{\partial L}{\partial v^A_\mu} (j^1 \phi)(\xi^A \circ \phi - \phi^{A,\nu} \xi^\nu) + L(j^1 \phi) \xi^\mu \right] d^{n+1}x \\
&= \left\{ \frac{\partial L}{\partial y^A}(j^1 \phi) - \frac{\partial}{\partial x^\mu} \left( \frac{\partial L}{\partial v^A_\mu}(j^1 \phi) \right) \right\} (-\xi^A \circ \phi + \phi^{A,\nu} \xi^\nu) \\
&\quad + \frac{\partial L}{\partial x^\mu}(j^1 \phi) \xi^\mu + \frac{\partial L}{\partial y^A}(j^1 \phi)(\xi^A \circ \phi) \\
&\quad + \frac{\partial L}{\partial v^A_\mu}(j^1 \phi) \left[ \frac{\partial \xi^A}{\partial x^\mu} + \frac{\partial \xi^A}{\partial y^B} \nu^B_\mu - \frac{\partial \xi^\nu}{\partial x^\mu} v^A_\nu \right](j^1 \phi) \\
&\quad + L(j^1 \phi) \frac{\partial \xi^\mu}{\partial x^\mu} \right\} d^{n+1}x \\
&= \left\{ \frac{\delta L}{\delta \phi^A}(\mathcal{L}_\xi \phi)^A + \delta \xi L \right\} (j^1 \phi) d^{n+1}x. \quad (4D.11)
\end{align*}
\]

Here \( \delta L/\delta \phi^A \) is the Euler–Lagrange derivative of \( \mathcal{L} \) defined in (3B.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 \( \mathfrak{g} \)-action is \textbf{vertically transitive} if for each \( x \in X \) and each \( y \in Y_x \) and each (local) section \( \phi \) through \( y \),

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

Looking forward to §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 \textit{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 4D.3.** Assume that the action of \( \mathfrak{g} \) on \( Y \) is vertically transitive and that the Lagrangian density is \( \mathfrak{g} \)-equivariant. Then a local section \( \phi \) 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 4D.4.** 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 $\phi$ with $\phi(x) = y$, the image of $\phi$ 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 $(x^\mu, y^A)$, 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 $\xi_Y$ 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-covariant) 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 $(\chi, \xi)$. Then (4D.8) reads

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

where $J^L(\xi) = (\partial L/\partial v^A)\xi^A$ 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 (3B.8) is Diff($\mathbb{R}$)-equivariant. Then the corresponding momentum map vanishes identically, consistent with the fact that $p = 0$, cf. (3B.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 (3B.12) is invariant under the prolongation of $\mathcal{F}(X)$ to $J^1(\Lambda^1 X)$. In this case (4D.10) reads

$$ (j^1 A)^* J^L(\chi) = (j^1 A)^* \left( F^{\nu \mu} \sqrt{-g} \chi_{,\nu} d^3 x_\mu \right) $$
$$ = (A^{\mu,\nu} - A^{\nu,\mu}) \sqrt{-g} \chi_{,\nu} d^3 x_\mu \quad (4D.17) $$

whence

$$ d \left[ (j^1 A)^* J^L(\chi) \right] = \left[ (A^{\mu,\nu} - A^{\nu,\mu}) \sqrt{-g} \right]_{,\mu} \chi_{,\nu} d^4 x $$

where we have used $dx^\lambda \wedge d^2 x_\mu = \delta^\lambda_\mu d^4 x$ and noted that the terms containing $\chi_{,\nu\mu}$ have dropped out by symmetry. Since $\chi$ is arbitrary, this quantity vanishes iff

$$ (A^{\mu,\nu} - A^{\nu,\mu})_{,\mu} = 0 \quad (4D.18) $$
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 is not $\text{Diff}(X)$-equivariant: the left hand side of the equivariance condition (4D.1), namely

$$L(\eta \cdot j^1 A) = L(j^1(\eta \cdot A)),$$

is not in general the same as its right hand side

$$\left(\eta^{-1}_{\chi}\right) L(j^1 A) = L(j^1 A) J^{-1},$$

$J$ being the Jacobian determinant of $\eta_{\chi}$. We can remedy this by allowing $g$ to vary parametrically, in which case (i) $F_{\mu\nu} F^{\mu\nu} = F_{\mu\nu} g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}$ will transform as a spacetime scalar, so that

$$L(\eta \cdot j^1(A; g)) = L(j^1(\eta \cdot A; \eta \cdot g)) = L(j^1(A; g)) J^{-1},$$

and (ii) the $\sqrt{-g}$ factor in $L$ will give rise to a factor of $J^{-1}$ which will balance the offending factor $J^{-1}$ in the above. Combined with our previous results, we see that parametrized electromagnetism is $(\text{Diff}(X) \ltimes \mathcal{F}(X))$-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 $\tilde{J}^\mathcal{L} = \mathcal{L}^* \tilde{J}$. Indeed, from (4D.11) we see that even if the section $A$ of $\Lambda^1X$ satisfies Maxwell’s equations,

$$d\left[j^1(A; g)^* \tilde{J}^\mathcal{L}(\xi, \chi)\right] = \frac{\partial L}{\partial g_{\sigma\rho}} (\mathcal{L} \xi g)_{\sigma\rho} d^4 x$$

$$= \mathcal{F}^\sigma_{\nu} \xi^\nu_{\sigma} d^4 x$$

need not vanish, where

$$\mathcal{F}^\sigma_{\rho} = - \left(F^\sigma_{\alpha\rho} F_{\alpha\beta} + \frac{1}{4} F_{\alpha\beta} F^\alpha_{\beta} \delta^\sigma_{\rho}\right) \sqrt{-g}$$

is the stress-energy-momentum tensor density 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 (3B.19) is equivariant under the action of $\text{Diff}(X)$, but not under that of $\mathcal{F}(X)$. Indeed, we compute
\[
\delta_{(\xi,\chi)} L = \frac{1}{2} \epsilon^{\mu\nu\sigma} F_{\mu\nu} \chi_{\sigma}.
\] (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
\[
(j^1 A)^* J^L (\xi, \chi) =
\left( \epsilon^{\mu\nu\sigma} (-A_{\tau, \tau} \xi^\tau_{, \nu} - A_{\nu, \tau} \xi^\tau + \chi_{, \nu}) + \epsilon^{\nu\tau\sigma} A_{[\tau, \nu]} \xi^\mu_{, \sigma} \right) A_\sigma d^2 x_{\mu},
\] (4D.21)
is conserved; we obtain
\[
\mathbf{d} \left[ (j^1 A)^* J^L (\xi, \chi) \right] = 2 \epsilon^{\mu\nu\sigma} A_{[\nu, \mu]} \left( A_{\sigma, \tau} \xi^\tau_{, \sigma} + A_{\tau, \sigma} \xi^\tau + \frac{1}{2} \chi_{, \sigma} \right)
\] which is zero when the Chern–Simons equations $A_{[\mu, \nu]} = 0$ 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 $\text{Diff}(X) \ltimes \mathcal{F}(X)$ 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 $A_{\nu, \mu}(x) - A_{\mu, \nu}(x) = 0$ at each $x \in X$.

**d Bosonic Strings.** The harmonic map Lagrangian density defined by equation (3B.24) is equivariant under the action of $\text{Diff}(X) \ltimes \mathcal{F}(X, \mathbb{R}^+)$ on $Y = (X \times M) \times X S^1$, using (4C.22)–(4C.24) to compute the Lie derivatives $\mathcal{L}_\xi \phi$, $\mathcal{L}_\xi h$, and $\mathcal{L}_\lambda h$ via (4D.9), (4D.10) gives
\[
j^1(\phi, h)^* J^L (\xi, \lambda) =
\sqrt{-h} g_{AB} \left( h^{\mu\nu} \phi^A_{, \rho} \phi^B_{, \nu} \xi^\rho - \frac{1}{2} h^{\rho\sigma} \phi^A_{, \sigma} \phi^B_{, \rho} \xi^\mu \right) d^1 x_{\mu}.
\] (4D.22)
Note that $\lambda$ does not appear in this expression. From this, we compute
\[
\mathbf{d} \left[ j^1(\phi, h)^* J^L (\xi, \lambda) \right] =
\left( \frac{\delta L}{\delta \phi^A_{, \nu}} \xi^\nu_{, \nu} + \frac{\delta L}{\delta h_{\sigma \rho}} (h_{\sigma, \rho} \xi^\nu + h_{\sigma, \nu} \xi^\rho + h_{\rho, \nu} \xi^\sigma) \right) d^2 x.
\] (4D.23)
At a given point $x \in X$, the quantities $\xi^\nu$ 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 (3B.32) as a Noether conservation law. The arbitrariness of $\xi^\nu$ then gives

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

which is not equivalent to the harmonic map equation (3B.31). This is consistent with the fact that $\mathcal{G}$ is not vertically transitive; indeed (4D.12) is not satisfied since, by (4C.22), $(\xi, \lambda)_Y$ has no component in the $\partial/\partial \phi^A$ direction.

We point out that if we had computed $d[ j^1(\phi, h)^* J^C(\xi, \lambda) ]$ directly from (4D.11), we would have obtained (4D.22) with the addition of the term

$$-2\lambda \frac{\delta L}{\delta h_{\alpha\beta}} h_{\alpha\beta}.$$ 

But this vanishes by virtue of (3B.33).
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 Chapter 8; for this discussion, we use the term loosely. It will be helpful in what follows to keep in mind the examples in Chapter 4, since the groups studied there are in fact the gauge groups of their respective theories.

Consider a classical field theory with gauge group $G$. Suppose that $G \subset \text{Aut}(Y)$, 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 $G$ and the (spacetime) diffeomorphism group $\text{Diff}(X)$. The first consists of those which are **parametrized** in the sense that the natural homomorphism $\text{Aut}(Y) \to \text{Diff}(X)$ given by $\eta_Y \mapsto \eta_X$ maps $G$ onto $\text{Diff}(X)$ (or at least onto a “sufficiently large” subgroup thereof, such as the compactly supported diffeomorphisms). This terminology—due to Kuchař [1973]—reflects the fact that such a theory is invariant under (essentially) arbitrary relabeling of the points (i.e., reparametrizations) of the “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 $G$ 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 topological field theories (see Horowitz [1989]).

At the other extreme, there are the **background** theories in which $G \subset \text{Aut}_{\text{id}}(Y)$, the automorphisms of $Y$ which cover the identity on $X$ (or at least in which the image of $G$ 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 $G$ 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. §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. §7D and Chapter 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 \( Y \to X \).

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 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 in which all fields are variational. 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 §11D. 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 (Figure I.1), indicating in which categories various field theories fall.

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 \( \phi : \Sigma \times \mathbb{R} \to M \)

---

10 Our terminology here is also at odds with the standard usage of “metric” and “non-metric” as applied to theories of gravity.
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 $\Sigma$ is the fluid “reference manifold” and $M$ is the physical 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 $\psi : X \to \Sigma \times \mathbb{R}$, 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) \text{ 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


Belinfante, F. J. [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.


REFERENCES


Dedecker, P. [1953], Calcul des variations, formes différentielles et champs géodésiques, *Colloq. Int. de Géométrie Différentielle* (Strasbourg), 17–34.


Hanson, A. J., T. Regge, and C. Teitelboim [1976], Constrained Hamiltonian systems, Accademia Nazionale dei Lincei (Rome) 22.


Kijowski, J. and W. Tulczyjew [1979], A symplectic framework for field theories, volume 107 of *Lecture Notes in Phys.*.


Künzle, H. P. and C. Duval [1986], Relativistic and non-relativistic classical field theory on five-dimensional spacetime, *Class. Quantum Grav.* 3, 957–973.


Moreau, J.-J. [1982], Fluid dynamics and the calculus of horizontal variations, 

Moser, J. and A. P. Veselov [1991], Discrete versions of some classical integrable 
systems and factorization of matrix polynomials, Comm. Math. Phys. 139, 

Muñoz, J. [1985], Poincaré–Cartan forms in higher order variational calculus on 
fibered manifolds, Rev. Mat. Iberoamericana 1, 85–126.

Nambu, Y. [1970], Duality and hydrodynamics, Lectures for the Copenhagen 
Summer Symposium, unpublished.

Neal M. O. and T. Belytschko [1989], Explicit-explicit subcycling with non-
integer time step ratios for structural dynamic systems. Computers & Struc-
tures, 6, 871–880.

Neishtadt, A. [1984], The separation of motions in systems with rapidly rotating 
phase, P. M. M. USSR 48, 133–139.

Novotny, J. [1982], On the geometric foundations of the Lagrange formulation of 
general relativity. In Soos, Gy. and J. Szenthe, editors, Differential Geometry, 

Olver, P. J. [1993], Applications of Lie Groups in Differential Equations, volume 

Ortega, J.-P. and T. S. Ratiu [2004], Momentum Maps and Hamiltonian Reduc-

Ouzilou, R. [1972], Expression symplectique des problèmes variationnels, Symp. 

Palais, R. [1968], Foundations of Global Nonlinear Analysis. Addison–Wesley, 
Reading, Massachusetts.

Pâquet, P. V. [1941], Les formes différentielles extérieurs $\Omega_n$ dans le calcul des 

Patrick, G. W. [1985], Singular Momentum Mappings, Presymplectic Dynamics 


