Momentum Maps and Classical Fields

Part III: Gauge Symmetries and Initial Value Constraints

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

*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.

Contents

III—Gauge Symmetries and Initial Value Constraints

This part studies gauge symmetries and the relation between various constraint sets and zero sets of instantaneous momentum and energy-momentum maps. In particular, the zero set $\mathcal{E}_\tau^{-1}(0)$ of the energy-momentum map for the gauge group $\mathcal{G}$ will be shown in Chapter 10 to coincide with the final constraint set $\mathcal{C}_\tau$. Similarly, the momentum map for a certain foliation $\dot{\mathcal{G}}_\tau$, which is derived from the gauge group $\mathcal{G}$ and is defined in Chapter 11, will be identified with the primary constraint set $\mathcal{P}_\tau$. Both these results require that the system under consideration be first class in an appropriate sense. We prepare for these results by defining the gauge group and discussing its properties in Chapter 8 and by proving the Vanishing Theorem in Chapter 9.

8 The Gauge Group

Here we define the gauge group of a given classical field theory, analyze its properties, and then show how to construct it. In the final section we discuss the correspondence between the notions of "gauge transformation" in the covariant and instantaneous formalisms.

8A Covariance, Localizability, and Gauge Groups

Suppose we have a group $\mathcal{G}$, as in §4D, which acts on Y by bundle automorphisms, so that we have a homomorphism $\mathcal{G} \to \mathrm{Aut}(Y)$. (We will often blur the distinction between $\mathcal{G}$ and its image in $\mathrm{Aut}(Y)$.) Our fundamental assumption **A1** is that the Lagrangian density $\mathcal{L}$ is equivariant with respect to the induced actions of $\mathcal{G}$ on J^1Y and $\Lambda^{n+1}X$. If this is the case, we say that the Lagrangian field theory under consideration is $\mathcal{G}$-***covariant***.

Covariance under the action of a group $\mathcal{G}$ leads to the following important consequence. Recall that Sol denotes the set of all spacetime solutions of the Euler–Lagrange equations.

Proposition 8A.1. *Suppose that a Lagrangian field theory is $\mathcal{G}$-covariant. Then the induced action of $\mathcal{G}$ on $\mathcal{Y}$ stabilizes* $\mathrm{Sol} \subset \mathcal{Y}$.

In other words, if $\mathcal{L}$ is $\mathcal{G}$-equivariant and $\phi \in \mathcal{Y}$ is a solution of the Euler–Lagrange equations, then so is $\eta \cdot \phi$ for all $\eta \in \mathcal{G}$. Here, and throughout the rest of Part III, we suppose that all fields are variational.

Proof. Let $\phi \in \mathcal{Y}$ be a solution of the Euler–Lagrange equations so that, by Theorem 3B.1

$$(j^1\phi)^*(V \rfloor \mathbf{d}\Theta_{\mathcal{L}}) = 0 \tag{8A.1}$$

for all $V \in \mathfrak{X}(J^1Y)$. For $\eta \in \mathcal{G}$, (4A.5)yields

$$\begin{aligned} j^1(\eta \cdot \phi)^*(V \rfloor \mathbf{d}\Theta_{\mathcal{L}}) &= (\eta_{J^1Y} \circ j^1\phi \circ \eta_X^{-1})^*(V \rfloor \mathbf{d}\Theta_{\mathcal{L}}) \\ &= (\eta_X^{-1})^*(j^1\phi)^*\eta_{J^1Y}{}^*(V \rfloor \mathbf{d}\Theta_{\mathcal{L}}). \end{aligned}$$

Since $\eta_{J^1(Y)}$ is a diffeomorphism, each vector field V on $J^1(Y)$ can be written in the form $V = T\eta_{J^1(Y)} \cdot W$ for some W. Substituting this into the preceding equation and using the fact that $\mathcal{G}$-equivariance of $\mathcal{L}$ implies $\Theta_{\mathcal{L}}$ is $\mathcal{G}$-invariant (cf. Proposition 4D.1), we obtain

$$\begin{aligned} j^1(\eta \cdot \phi)^*(V \rfloor \mathbf{d}\Theta_{\mathcal{L}}) &= (\eta_X^{-1})^*(j^1\phi)^*(W \rfloor \eta_{J^1Y}{}^*\mathbf{d}\Theta_{\mathcal{L}}) \\ &= (\eta_X^{-1})^*(j^1\phi)^*(W \rfloor \mathbf{d}\Theta_{\mathcal{L}}) = 0 \end{aligned}$$

by (8A.1). The result now follows from Theorem 3B.1 ■

If a field theory is $\mathcal{G}$-covariant, then, according to this proposition, $\mathcal{G}$ acts by ***symmetries*** in the sense that it maps solutions of the Euler–Lagrange equations to solutions. A very important property of a *gauge* group, which distinguishes it from a mere symmetry group, is that the former is "localizable."

We say that a group $\mathcal{G}$ of automorphisms of a bundle $Y \to X$ is ***localizable*** provided that for each pair of disjoint hypersurfaces Σ_1 and Σ_2 in X and each Lie algebra element $\xi \in \mathfrak{g}$, there is a Lie algebra element $\chi \in \mathfrak{g}$ such that $\chi_Y | \pi_{XY}^{-1}(\Sigma_1) = \xi_Y | \pi_{XY}^{-1}(\Sigma_1)$ and $\chi_Y | \pi_{XY}^{-1}(\Sigma_2) = 0$.

Now consider a Lagrangian field theory and a subgroup $\mathcal{G}$ of $\mathrm{Aut}(Y)$. We say that $\mathcal{G}$ is a ***gauge group*** for the theory provided $\mathcal{G}$ is localizable and the theory is $\mathcal{G}$-covariant. We will justify this terminology and discuss our standard examples in the next section.

A constrained theory may have no gauge symmetries at all (that is, $\mathcal{G}$ is the zero group); the Proca field on a fixed background spacetime is such a theory. (See Gotay and Nester [1980]). But the Proca field acquires the gauge group $\mathrm{Diff}(X)$ when parametrized. More generally, every parametrized theory has nontrivial gauge symmetry. At the other extreme, there are systems which are totally gauge in the sense that they have no true dynamical degrees of freedom; for instance, parametrized electromagnetism on a $(1+1)$-dimensional spacetime (cf. Example 6E.**b**).

8B Principal Bundle Construction of the Gauge Group

The gauge groups of many field theories arise via the following principal bundle construction (cf. Fischer [1982]). One can often associate a principal bundle B to the configuration bundle Y in such a way that automorphisms of B induce automorphisms of Y; the association may, but need not be, via the standard notion of an associated bundle.

Let $\pi_{XB} : B \to X$ be a principal bundle with Lie group G. (Thus, G acts effectively on B on the right, leaving B fiberwise invariant.) Suppose that $\mathcal{G} \subset \mathrm{Aut}(B)$ is a group of automorphisms of B, i.e., fiber-preserving diffeomorphisms of B which commute with the G-action. In general we have a homomorphism $\mathrm{Aut}(B) \to \mathrm{Diff}(X)$ which maps $\eta \in \mathrm{Aut}(B)$ to the induced diffeomorphism η_X of X. We assume that Y is "associated" to B in the sense that $\mathcal{G}$ acts on Y by bundle automorphisms that cover the induced action on X. In most examples, the choices of B and $\mathcal{G}$ are usually apparent.

Examples

We refer the reader to §4D for a discussion of $\mathcal{G}$-covariance in each case.

a Particle Mechanics. In parametrized particle dynamics, B is the identity bundle $\mathbb{R} \to \mathbb{R}$ (so that $G = \{\mathrm{Id}\}$), $\mathcal{G} = \mathrm{Aut}(B) = \mathrm{Diff}(\mathbb{R})$ and Y is the associated bundle $B \times_G Q = \mathbb{R} \times Q$ with typical fiber Q.

Obviously the group $\mathrm{Diff}(\mathbb{R})$ of time reparametrizations is localizable. Thus $\mathrm{Diff}(\mathbb{R})$ is a gauge group.

b Electromagnetism. For electromagnetism as a 1-form field theory on a fixed background X, G is the additive group $\mathbb{R}$ and $B = X \times \mathbb{R}$. Then we have $\mathcal{G} = \mathrm{Aut}_{\mathrm{Id}}(B) = \mathcal{F}(X)$, where $\mathrm{Aut}_{\mathrm{Id}}(B)$ is the group of automorphisms covering the identity on X. In this case Y is the generalized associated bundle

$$J^1(X \times \mathbb{R})/\mathbb{R} \approx \Lambda^1 X$$

and the induced action of $\mathcal{G}$ on Y coincides with (4C.17).

The group $\mathcal{G} = \mathcal{F}(X)$ is manifestly localizable, and so it is a gauge group.

To illustrate how special gauge groups are, as compared to symmetry groups, we consider other possible choices of $\mathcal{G}$ in the context of this example. For instance, take $\mathcal{G}$ to be the additive group of closed 1-forms on X, acting on Y by $(\alpha, A) \mapsto A + \alpha$. Then $\mathcal{G}$ is a symmetry group of $\mathcal{L}$, but it need not be localizable. To ensure localizability, we need to restrict to the additive group of *exact* 1-forms on X, i.e., $\mathcal{G} \approx \mathcal{F}(X)$. Now suppose we take $\mathcal{G}$ to be the isometry group of the fixed background spacetime X, acting on $\Lambda^1 X$ by push-forward. Then electromagnetism is $\mathcal{G}$-covariant, but $\mathcal{G}$ is not a gauge group as again it is not localizable. On the other hand, $\mathcal{G} = \mathrm{Diff}(X)$ is localizable but the Maxwell Lagrangian density (3B.12), is not $\mathrm{Diff}(X)$-covariant *unless* gravity is included, either parametrically or dynamically. Similar remarks obviously apply to any theory on a background spacetime.

When the spacetime metric is not fixed, we keep $B = X \times \mathbb{R}$ as in the background case. However, since in this context the metric on X is supposed to transform under spacetime diffeomorphisms, we now take

$$\tilde{\mathcal{G}} = \mathrm{Aut}(B) = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$$

to be the full automorphism group of B. Then $\tilde{\mathcal{G}}$ is localizable and acts on the "parametrized" configuration bundle $\tilde{Y} = \Lambda^1 X \times_X S_2^{3,1}(X)$ as in Example **b** of §4C.

c A Topological Field Theory. Again we take G to be the additive group $\mathbb{R}$ and $B = X \times \mathbb{R}$. Then $\mathcal{G} = \mathrm{Aut}(B) = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$ where, as before, $Y = J^1(X \times \mathbb{R})/\mathbb{R} \approx \Lambda^1 X$ and the induced action of $\mathcal{G}$ on Y is given by (4C.17) (4C.18).

Although the Chern–Simons theory is not $\mathrm{Diff}(X) \ltimes \mathcal{F}(X)$-covariant, a glance back at the proof of Proposition 8A.1 shows that it remains valid provided merely that $\Omega_{\mathcal{L}} = -\mathbf{d}\Theta_{\mathcal{L}}$ is invariant, which it is in this instance.

d Bosonic Strings. This is similar to parametrized electromagnetism, cf. Example **b** above. Keeping the conformal invariance of the bosonic string in mind, we let G be the multiplicative group $\mathbb{R}^+$ and set $B = X \times \mathbb{R}^+$. Then

$$\mathcal{G} = \mathrm{Aut}(X \times \mathbb{R}^+) = \mathrm{Diff}(X) \ltimes \mathcal{F}(X, \mathbb{R}^+)$$

acts on the configuration bundle Y for bosonic strings as in Example **d** of §4C.

Remark 8B.1. When a field theory is parametrized, the above examples show that the appropriate principal bundle B is just that corresponding to the "internal" degrees of freedom. It follows that the appropriate group $\mathcal{G} = \mathrm{Aut}(B)$ has $\mathrm{Diff}(X)$ built in automatically. (We refer the reader to Interlude I for a discussion of parametrized vs. background theories.)

When the fields being studied propagate on a fixed background spacetime X, we may proceed in two ways. In one approach, we restrict $\mathcal{G}$ to be just $\mathrm{Aut}_{\mathrm{Id}}(B)$—those automorphisms of B which restrict to the identity on X. Hence $\mathcal{G}$ corresponds to the "internal" gauge group. This approach leads to the appropriate relationship between the energy-momentum map and the constraints for these fields (see Chapters 10 and 11). In the second approach, we let $\mathcal{G}$ be all of $\mathrm{Aut}(B)$ so that $\mathrm{Diff}(X)$ is now "included" and the background metric is treated parametrically. This approach, according to §7D, leads to the desired relationship between the energy-momentum map and the Hamiltonian. Example **b** shows that only when the metric on X is dynamic do we get both desired relations.

Thus, on a group-theoretical level the parametrization of a theory can be accomplished merely by replacing $\mathrm{Aut}_{\mathrm{Id}}(B)$ by $\mathrm{Aut}(B)$. ♦

8C Gauge Transformations

In §6E we defined a ***gauge transformation*** in the instantaneous formalism to be a diffeomorphism of the final constraint set $\mathcal{C}_\tau$ which stabilizes the fibers of the subbundle $T\mathcal{C}_\tau \cap T\mathcal{C}_\tau^\perp$ of $T\mathcal{C}_\tau$ (or, what essentially amounts to the same thing, a flow on $\mathcal{C}_\tau$ whose generating vector field belongs to the distribution $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$). In this section we relate this notion of gauge transformation to its covariant counterpart, defined in §8A, and discuss what it means for the gauge group to be "full."

Consider a subgroup $\mathcal{G}$ of $\mathrm{Aut}(Y)$. The first order of business is to determine how $\mathcal{G}$ "acts" in the instantaneous framework. Strictly speaking, of course, $\mathcal{G}$ need not act at all in this context, as it does not necessarily stabilize Cauchy surfaces. Nonetheless, using the instantaneous energy-momentum map $\mathcal{E}_\tau : \mathcal{P}_\tau \to \mathfrak{g}^*$ defined in §7D, we show that $\mathfrak{g}$ naturally defines a Lie subalgebra $\mathfrak{g}_{\mathcal{C}_\tau}$ of $\mathfrak{X}(\mathcal{C}_\tau)$ which, for our purposes, serves as an (infinitesimal) action.

Proposition 8C.1. *Consider a $\mathcal{G}$-covariant field theory. For every $\xi \in \mathfrak{g}$, there exist solutions $\xi_{\mathcal{C}_\tau} \in \mathfrak{X}(\mathcal{C}_\tau)$ of the equation*

$$(\xi_{\mathcal{C}_\tau} \mathbin{\lrcorner} \omega_\tau - \mathbf{d}\langle \mathcal{E}_\tau, \xi\rangle) \,|\, \mathcal{C}_\tau = 0. \tag{8C.1}$$

The proof of this Proposition is deferred until §10A, as it relies upon certain results to be established there. Below we prove a modified version with $\mathfrak{g}$ replaced by $\mathfrak{g}_\tau$, the stabilizer algebra of the Cauchy surface Σ_τ.

Proposition 8.2 says that, for each $\xi \in \mathfrak{g}$, $\langle \mathcal{E}_\tau, \xi\rangle$ has Hamiltonian vector fields along $\mathcal{C}_\tau$ which are tangent to $\mathcal{C}_\tau$. Given ξ, (8C.1) only defines $\xi_{\mathcal{C}_\tau}$ modulo elements of $\ker \omega_\tau \cap \mathfrak{X}(\mathcal{C}_\tau)$. Denote by $\mathfrak{g}_{\mathcal{C}_\tau}$ the involutive distribution generated by the set $\{\xi_{\mathcal{C}_\tau} \mid \xi \in \mathfrak{g}\}$. We may think of $\mathfrak{g}_{\mathcal{C}_\tau} \subset \mathfrak{X}(\mathcal{C}_\tau)$ as playing the role of an "action" of $\mathfrak{g}$ on $\mathcal{C}_\tau$. A main result is:

Theorem 8C.2. *Assume that all fields are variational. For each gauge group $\mathcal{G}$,*

$$\mathfrak{g}_{\mathcal{C}_\tau} \subset \mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp.$$

Again, the proof will be postponed until §10A. The standing hypothesis that all fields are variational is crucial; see Example **b** following for an illustration of what can happen when this is not the case.

Recall from §6E that the distribution $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$ is the gauge algebra of the theory. Thus this Theorem constitutes the fundamental link between the covariant and instantaneous notions of "gauge." It follows that *if $\mathcal{G}$ is a gauge group in the covariant sense and all fields are variational, then $\mathcal{G}$ "acts" on each $\mathcal{C}_\tau$ by gauge transformations in the instantaneous sense.*

To ensure that $\mathcal{G}$ comprises *all* the gauge freedom of the theory, we shall require

A5 Fullness. $\mathfrak{g}_{\mathcal{C}_\tau} = \mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$.

A gauge group $\mathcal{G}$ is called ***full*** if this condition holds.

This condition should be thought of as a means of checking whether the choice of gauge group is "correct." We emphasize that this is a nontrivial matter in general. Our philosophy throughout this work is that one knows the correct gauge group at the outset; this is indeed often the case, as in the examples we have presented. But from a practical standpoint, for a given Lagrangian field theory, the choice of $\mathcal{G}$ may not be entirely obvious. Thus one may "undershoot," that is, choose a candidate gauge group which is too small. This error would be signaled by a violation of **A5**. We illustrate the sort of things that can go wrong when the gauge group is not full in Chapter 10. We point out, however, that it is not possible to "overshoot," at least when all fields are variational. This is because of Theorem 8C.2; a candidate gauge group which is too large (in the sense that $\mathfrak{g}_{\mathcal{C}_\tau} \supset \mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$) would have to violate either localizability or covariance, and so could be immediately discarded.

Remark 8C.3. Note that by definition, $\mathfrak{g}_{\mathcal{C}_\tau}$ *always* contains $\ker \omega_\tau \cap \mathfrak{X}(\mathcal{C}_\tau)$ as an involutive subdistribution. Consequently, if a given theory has only *primary* constraints, so that $\mathcal{C}_\tau = \mathcal{P}_\tau$, then any gauge group is vacuously full, since in this case $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp = \ker \omega_\tau$. This is true even for the zero group. This observation and the closing comments in §6E indicate that the fullness of $\mathcal{G}$ is correlated with first class *secondary* constraints, but not with first class *primaries*. The connection between $\mathcal{G}$ and the first class secondaries will be made more precise in Chapter 10. We also remark that there is a condition on $\mathcal{G}$, similar to fullness, which is explicitly tailored to the first class primary constraints; see §11D. ♦

Remark 8C.4. A computable sufficient condition for fullness is given in Corollary 10B.6. ♦

Remark 8C.5. It is conceivable that a given theory could have *several* different full gauge groups. (See Example **d** following for an illustration, as well as the footnote to Example 4C.**b**.) However, up to questions of reconstructing groups from algebras[25] and the faithfulness of the representation $\xi \mapsto \xi_{\mathcal{C}_\tau}$, assumption **A5** makes the gauge group "unique," simply because $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$ is intrinsic to the theory, independent of any choice of gauge group.[26] ♦

We now state and prove a modified version of Proposition 8C.1, as promised earlier. Recall from Section 7C that the subgroup $\mathcal{G}_\tau$ of $\mathcal{G}$ *does* act in the instantaneous formalism with momentum map $\mathcal{J}_\tau : T^*\mathcal{Y}_\tau \to \mathfrak{g}_\tau^*$.

Proposition 8C.6. *Assume* **A1**, **A2**, *and* **A4**. *Then the action of* $\mathcal{G}_\tau$ *stabilizes both the instantaneous primary and final constraint sets* $\mathcal{P}_\tau$ *and* $\mathcal{C}_\tau$ *in* $T^*\mathcal{Y}_\tau$.

Proof. By Proposition 4D.1 the covariant Legendre transform is equivariant with respect to the actions of $\mathcal{G}_\tau$ on $(J^1Y)_\tau$ and Z_τ, so that $\mathcal{G}_\tau$ stabilizes N_τ

[25] We will not consider global issues—such as the existence of 'large' gauge transformations (viz., the connectedness of $\mathcal{G}$)—in this book.

[26] More precisely, it makes the action on $\mathcal{C}_\tau$ unique.

in Z_τ and hence $\mathcal{N}_\tau$ in $\mathcal{Z}_\tau$. Then (7C.2) and Corollary 6C.5 imply that the $\mathcal{G}_\tau$-action stabilizes $\mathcal{P}_\tau$ in $T^*\mathcal{Y}_\tau$.

Recall from (6D.10) that the canonical decomposition map $\mathrm{can}_\tau : \mathcal{Y} \to \mathcal{P}_\tau$ relative to the embedding $\tau : \Sigma \to X$ is given by

$$\mathrm{can}_\tau(\phi) = R_\tau(\mathbb{F}\mathcal{L} \circ j^1\phi \circ \tau).$$

The above results yield, for all $\eta \in \mathcal{G}_\tau$,

$$\mathrm{can}_\tau(\eta_{\mathcal{Y}}(\phi)) = \eta_{T^*\mathcal{Y}_\tau}(\mathrm{can}_\tau(\phi)).$$

In view of **A2** and **A4**, Corollary 6E.11 asserts that $\mathrm{can}_\tau(\mathrm{Sol}) = \mathcal{C}_\tau$; to prove that $\mathcal{G}_\tau$ stabilizes $\mathcal{C}_\tau$, it therefore suffices to show that $\mathcal{G}_\tau$ stabilizes Sol in $\mathcal{Y}$. But this is immediate from our assumption of covariance together with Proposition 8A.1. ■

Now, on $T^*\mathcal{Y}_\tau$ we may solve

$$\xi_{T^*\mathcal{Y}_\tau} \lrcorner\, \omega_{T^*\mathcal{Y}_\tau} = \mathbf{d}\langle \mathcal{J}_\tau, \xi\rangle \tag{8C.2}$$

uniquely for the generators $\xi_{T^*\mathcal{Y}_\tau}$. Finally, Proposition 8C.6 together with Corollary 7D.2(*ii*) guarantee that (8C.2), when pulled back to $\mathcal{P}_\tau$, coincides with (8C.1) for $\xi \in \mathfrak{g}_\tau$, and that $\xi_{T^*\mathcal{Y}_\tau}|\mathcal{C}_\tau$ is tangent to $\mathcal{C}\tau$.

Examples

a Particle Mechanics. Fix a Cauchy "surface" Σ_t. We know from Example **a** of §6E that $\mathcal{C}_t = \mathcal{P}_t$, and from Example **a** of §7D that $\mathcal{E}_t \equiv 0$. Thus (8C.1) yields $\mathfrak{g}_{\mathcal{C}_t} = \ker \omega_t$, so the gauge group $\mathrm{Diff}(\mathbb{R})$ is vacuously full, in accordance with Remark 8C.3.

b Electromagnetism. In the background case $\mathcal{G} = \mathcal{G}_\tau$, so we obtain a genuine action of $\mathcal{G} = \mathcal{F}(X)$ on $T^*\mathcal{Y}_\tau$. It is given by

$$(f, (A, \mathfrak{E})) \mapsto (A + \mathbf{d}f\,|\,\Sigma_\tau, \mathfrak{E})$$

for $f \in \mathcal{F}(X)$. For $\chi \in \mathfrak{g} \approx \mathcal{F}(X)$, we may write

$$\chi_{T^*\mathcal{Y}_\tau} = \chi_0 \frac{\delta}{\delta A_0} + \nabla\chi_\tau \cdot \frac{\delta}{\delta \mathbf{A}}$$

in adapted coordinates, where $\chi_0 := \chi_{,0}\,|\,\Sigma_\tau$ and $\chi_\tau = \chi\,|\,\Sigma_\tau$. Now χ_0 and χ_τ are independently specifiable along Σ_τ and so, by referring to Example **b** of §6E, we see that **A5** is satisfied. Thus $\mathcal{F}(X)$ is the full gauge group for background electromagnetism.

For parametrized electromagnetism it is no longer true that $\tilde{\mathcal{G}} = \tilde{\mathcal{G}}_\tau$, so we must compute the Hamiltonian vector fields directly. Given $(\xi, \chi) \in \tilde{\mathfrak{g}} \approx \mathfrak{X}(X) \ltimes \mathcal{F}(X)$, from (6C.22) and (7D.12) we eventually obtain:

$$(\xi,\chi)_{\mathcal{C}_\tau} = -\left(\xi^0 N\gamma^{-1/2}\gamma_{ij}\mathfrak{E}^j + \frac{1}{N\sqrt{\gamma}}(\xi^0 M^j + \xi^j)\mathfrak{F}_{ji} + D_i(\xi^\mu A_\mu - \chi)\right)\frac{\delta}{\delta A_i}$$
$$- D_j\left(\xi^0\gamma^{ik}\gamma^{jm}\mathfrak{F}_{km} + \left[(\xi^0 M^i + \xi^i)\mathfrak{E}^j - (\xi^0 M^j + \xi^j)\mathfrak{E}^i\right]\right)\frac{\delta}{\delta\mathfrak{E}^i}$$

modulo terms in the direction $\delta/\delta A_0$. Because the metric on X is not variational Theorem 8C.2 fails and these vector fields need not lie in $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$. But they do when $\xi^\mu = 0$, and the same argument as in the background case shows that $\tilde{\mathcal{G}} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$ is “overfull” in the sense that $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp \subset \mathfrak{g}_{\mathcal{C}_\tau}$.

c A Topological Field Theory. From (6C.30) and (7D.14) we compute

$$(\xi,\chi)_{\mathcal{C}_\tau} = D_i(\xi^\mu A_\mu - \chi)\left(\epsilon^{0ij}\frac{\delta}{\delta\pi^j} - \frac{\delta}{\delta A_i}\right)$$

modulo the directions $\delta/\delta A_0$. Referring back to §6E, it follows from this and the argument given in Example **b** above that $\mathcal{G} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$ is full. Note also that $\mathrm{Diff}(X)$ alone is *not* full.

This example indicates that it may suffice—and indeed may be necessary—to require just that the *action* of a field theory, as opposed to the Lagrangian density itself, be invariant for the main results of this chapter to hold.

d Bosonic Strings. Consider $(\xi, \lambda) \in \mathfrak{g} \approx \mathfrak{X}(X) \ltimes \mathcal{F}(X)$. From (7D.17) and (6C.37) we compute

$$(\xi,\lambda)_{\mathcal{C}_\tau} = \left(\frac{\xi_0 N}{2\sqrt{\gamma}}g^{AB}\pi_B + \xi^1 M D\varphi^A\right)\frac{\delta}{\delta\varphi^A}$$
$$+\left(\frac{\xi^0}{2\sqrt{\gamma}}g_{AB}D(ND\varphi^B) + \xi^1 D(M\pi_A)\right)\frac{\delta}{\delta\pi_A}$$

modulo the directions $\delta/\delta h_{\sigma\rho}$. In view of (6E.31), this may be rewritten

$$(\xi,\lambda)_{\mathcal{C}_\tau} = \xi^0 X_{N\mathfrak{H}} + \xi^1 X_{M\mathfrak{J}}.$$

Since ξ^0 and ξ^1 are independently specifiable, it follows from the computation of $\mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$ in Example **d** of §6E that the gauge group $\mathrm{Diff}(X) \ltimes \mathcal{F}(X, +)$ is full in this case. The subgroup $\mathrm{Diff}(X)$ is full as well.

9 The Vanishing Theorem and Its Converse

Here we begin the program of relating initial value constraints to zero levels of the energy-momentum map by proving the Vanishing Theorem (also called the second Noether theorem). This says that when evaluated on any solution of the Euler–Lagrange equations, the covariant momentum map integrated over a hypersurface is zero. The result is a consequence of Noether's theorem and localizablility.

9A The Vanishing Theorem

Let Σ_+ and Σ_- denote two hypersurfaces in X which form the boundary of a compact region; we call Σ_+ and Σ_- an ***admissible pair*** of hypersurfaces. (We have in mind the case of a spacetime X with Σ_+ and Σ_- deformations of a given hypersurface to the future and past, respectively, as in Figure 9.1.)

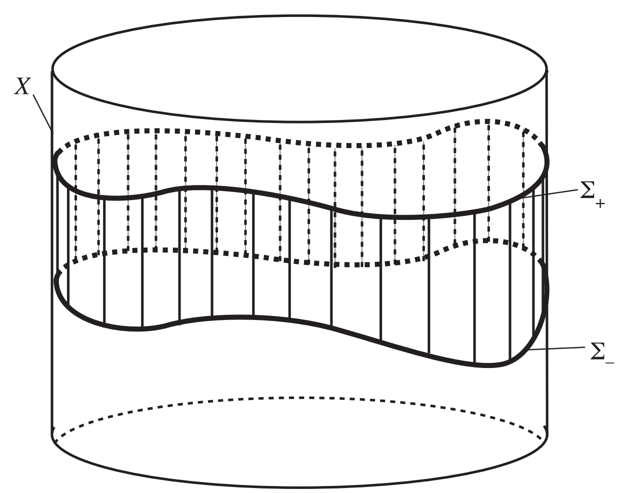


Figure 9.1: An admissible pair of hypersurfaces

Lemma 9A.1. *For every compact oriented hypersurface Σ there is a disjoint hypersurface Σ' such that (Σ, Σ') form an admissible pair.*

Proof. Choose an outward pointing normal vector field along Σ, and smoothly extend it to a vector field n on a neighborhood U of Σ in X. Since Σ is compact, there is a neighborhood $V \subset U$ of Σ on which the flow F of n is defined on some interval $(-k, k)$. Now take $\Sigma' = F_{k/2}(\Sigma)$; then $\Sigma \cup \Sigma'$ is the boundary of the compact region $F([0, k/2] \times \Sigma)$ in X. ■

Suppose that $\mathcal{G}$ is a localizable group of automorphisms of Y. Then for every compact oriented hypersurface Σ and every $\xi \in \mathfrak{g}$, there exists a disjoint hypersurface Σ' and a $\xi' \in \mathfrak{g}$ such that (Σ, Σ') form an admissible pair and

$$\xi'_Y \,|\, Y_\Sigma = \xi_Y \,|\, Y_\Sigma \quad \text{and} \quad \xi'_Y \,|\, Y'_\Sigma = 0,$$

where $Y_\Sigma = \pi_{XY}^{-1}(\Sigma)$ and $Y'_\Sigma = \pi_{XY}^{-1}(\Sigma')$. Roughly speaking, this means that we can localize any $\xi \in \mathfrak{g}$ by "shutting it off" on a hypersurface to the "past" or "future."

Now suppose that we have a field theory with gauge group $\mathcal{G}$ in which all fields are variational. Then from the first Noether theorem (Theorem 4D.2) and Stokes' theorem, we obtain

$$\int_{\Sigma_+} \tau_+^*(j^1\phi)^* J^{\mathcal{L}}(\xi) = \int_{\Sigma_-} \tau_-^*(j^1\phi)^* J^{\mathcal{L}}(\xi). \tag{9A.1}$$

for every solution ϕ of the Euler–Lagrange equations and admissible pair of hypersurfaces Σ_+ and Σ_-, where $\tau_\pm : \Sigma_\pm \to X$ are the inclusions. Note that, as always, the compactness assumption in this discussion can be relaxed provided all fields fall off sufficiently rapidly at "infinity."

Localizability and (9A.1) immediately lead to the "second Noether theorem":

Theorem 9A.2. (Vanishing Theorem) *Let $\mathcal{L}$ be the Lagrangian density for a field theory with gauge group $\mathcal{G}$. Then for any solution ϕ of the Euler–Lagrange equations and hypersurface Σ, the energy-momentum map on Σ in the Lagrangian representation vanishes:*

$$\int_{\Sigma} \tau^*(j^1\phi)^* J^{\mathcal{L}}(\xi) = 0$$

for all $\xi \in \mathfrak{g}$, where $\tau : \Sigma \to X$ is the inclusion.

9B The Converse of the Vanishing Theorem

This will be obtained using the results of §4D.

Theorem 9B.1. (Converse of the Vanishing Theorem) *Suppose $\mathfrak{L}$ is equivariant with respect to a vertically transitive group action. Let ϕ be a section of Y which satisfies*

$$\int_\Sigma \tau^*(j^1\phi)^* J^{\mathfrak{L}}(\xi) = 0 \tag{9B.1}$$

for all hypersurfaces Σ, where $\tau : \Sigma \to X$ is the inclusion. Then ϕ is a solution to the Euler–Lagrange equations.

Proof. Let Σ_+ and Σ_- be the images of two such embeddings τ_+ and τ_- which enclose a "bubble" U as in Figure 9.2. Then

$$\int_U \mathbf{d}\left[(j^1\phi)^* J^{\mathfrak{L}}(\xi)\right] = \int_{\Sigma_+} \tau_+^*(j^1\phi)^* J^{\mathfrak{L}}(\xi) - \int_{\Sigma_-} \tau_-^*(j^1\phi)^* J^{\mathfrak{L}}(\xi) = 0.$$

The open sets of the form U comprise a neighborhood base for X. Therefore

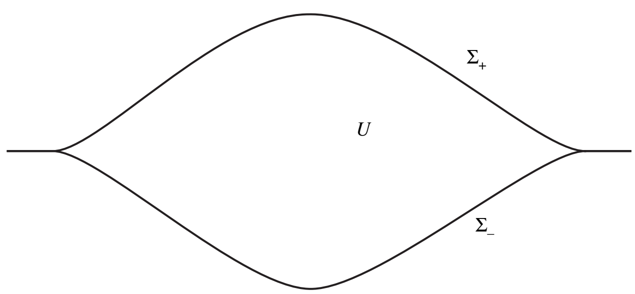


Figure 9.2: A bubble in spacetime

$\mathbf{d}\left[(j^1\phi)^* J^{\mathfrak{L}}(\xi)\right] = 0$ and so by Theorem 4D.3, ϕ is a solution to the Euler–Lagrange equations. ■

Theorem 9B.1 can be viewed as one formulation of the "geometrodynamics regained" program of Kuchař [1974]. Indeed, it essentially states that from symmetry assumptions and the vanishing of the energy-momentum map it is possible to recover the Euler–Lagrange equations; that is, *the geometry and symmetry together determine the field theory.*

Examples

In each of the following examples, we compute the energy-momentum map in the Lagrangian representation:

$$\int_\Sigma \tau^*(j^1\phi)^* J^{\mathfrak{L}}(\xi)$$

without assuming that ϕ is a solution of the Euler–Lagrange equations. Setting this equal to zero for all values of ξ, we find that for fixed Σ equation (9B.1) is equivalent to a subset of the Euler–Lagrange equations along Σ.

a Particle Mechanics. For time reparametrization-covariant particle dynamics with no internal symmetries, (4D.15) yields

$$(j^1\phi)^* J^{\mathfrak{L}}(f) = 0$$

as a matter of course, cf. Example **a** in §4D. Thus the Vanishing Theorem just reproduces this fact in this case. It gives no information at all regarding the equations of motion, since the action of $\mathcal{G} = \mathrm{Diff}(\mathbb{R})$ is vertically trivial.

b Electromagnetism. From (4C.12) we have

$$(j^1 A)^* J^{\mathfrak{L}}(\chi) = \mathfrak{F}^{\nu\mu}\chi_{,\nu} d^3x_\mu.$$

If we take Σ to be the Cauchy surface $x^0 = 0$ then, by virtue of the antisymmetry of $\mathfrak{F}^{\nu\mu}$, the left hand side of (9B.1) reduces to

$$\int_\Sigma \mathfrak{F}^{i0}\chi_{,i} d^3x_0.$$

Setting this equal to zero for all $\chi \in \mathcal{F}(X)$, and integrating by parts leads to Gauss' Law ${\mathfrak{F}^{i0}}_{,i} = 0$. On the other hand, if Σ is given by $x^k = 0$, we obtain

$$\int_\Sigma \mathfrak{F}^{\nu k}\chi_{,\nu} d^3x_k \qquad \text{(no sum on } k\text{)}.$$

Using the fact that $\mathfrak{F}^{kk} = 0$ (no sum), we may again integrate by parts, and so the vanishing of this expression for all $\chi \in \mathcal{F}(X)$ gives the rest of Maxwell's equations ${\mathfrak{F}^{\nu k}}_{,\nu} = 0$. Thus we get all of Maxwell's equations, consistent with the Converse of the Vanishing Theorem and the vertical transitivity of $\mathcal{G} = \mathcal{F}(X)$.

c A Topological Field Theory. From (4D.21), the energy-momentum map is

$$\int_\Sigma \Big(\epsilon^{\mu\nu\sigma}(-A_\tau \xi^\tau{}_{,\nu} - A_{\nu,\tau}\xi^\tau + \chi_{,\nu}) + \frac{1}{2}\epsilon^{\nu\tau\sigma}F_{\nu\tau}\xi^\mu\Big)A_\sigma \, d^2x_\mu.$$

Taking Σ to be the $x^0 = 0$ surface and integrating by parts, the terms involving ξ drop out and this reduces to

$$-\int_\Sigma \epsilon^{0ij}\chi A_{j,i}\, d^2x_0.$$

Upon requiring this to equal zero, the arbitrariness of χ implies that $F_{12} = 0$. Similarly, taking Σ to be the $x^k = 0$ surface for $k = 1, 2$ we obtain $F_{0k} = 0$. Thus the conclusion of Theorem 9B.1 holds even though the Chern–Simons theory is not $\mathcal{F}(X)$-invariant.

d Bosonic Strings. From (4D.8), (3B.24), and (4C.22)–(4C.24), we compute

$$\begin{aligned} j^1(\phi,h)^* J^{\mathcal{L}}(\xi,\lambda) = \sqrt{|h|}g_{AB}\Big[\Big\{&\frac{1}{2}(h^{00}\phi^A{}_{,0}\phi^B{}_{,0} - h^{11}\phi^A{}_{,1}\phi^B{}_{,1})\xi^0 \\ &+ h^{0\mu}\phi^B{}_{,\mu}\phi^A{}_{,1}\xi^1\Big\}d^1x_0 \\ &+ \Big\{\frac{1}{2}(h^{11}\phi^A{}_{,1}\phi^B{}_{,1} - h^{00}\phi^A{}_{,0}\phi^B{}_{,0})\xi^1 \\ &+ h^{1\mu}\phi^B{}_{,\mu}\phi^A{}_{,0}\xi^0\Big\}d^1x_1\Big]. \end{aligned}$$

First suppose that Σ is the Cauchy curve $x^0 = 0$. Then if (9B.1) is to hold for all (ξ, λ), we must have

$$g_{AB}(h^{00}\phi^A{}_{,0}\phi^B{}_{,0} - h^{11}\phi^A{}_{,1}\phi^B{}_{,1}) = 0 \qquad (9B.2)$$

and

$$g_{AB}h^{0\mu}\phi^B{}_{,\mu}\phi^A{}_{,1} = 0. \qquad (9B.3)$$

Using (3B.25)–(3B.27)and (5C.10), one directly verifies that (9B.3) is just the supermomentum constraint (6E.30), while (9B.2) is a combination of the supermomentum constraint and the superhamiltonian constraint (6E.29).

Next suppose that Σ is the timelike curve $x^1 = 0$. Then the vanishing of (9B.1) leads to (9B.2) again and

$$g_{AB}h^{1\mu}\phi^B{}_{,\mu}\phi^A{}_{,0} = 0. \qquad (9B.4)$$

Equations (9B.2)–(9B.4) are not independent; in fact

$$h^{00}(9B.4) - h^{11}(9B.3) = h^{01}(9B.2).$$

Together these equations are equivalent to the conformal Euler-Lagrange equation (3B.32). (Recall that of these latter three equations one is an identity.) Observe that we do not obtain the harmonic map equation (3B.31).—this is because $\mathcal{G}$ is not vertically transitive.

Thus while the Vanishing Theorem holds in this example and yields the conformal Euler-Lagrange equation (or, equivalently, the superhamiltonian and supermomentum constraints), its converse fails as it does not recover the harmonic map equation.

---

10 Secondary Constraints and the Instantaneous Energy-Momentum Map

The examples in the previous chapter illustrate the general result that when Σ is a Cauchy surface, the vanishing of the instantaneous energy-momentum map over Σ yields (first class secondary) initial value constraints. Our goals now are to prove this statement for a wide variety of field theories and to show under some reasonable hypotheses that the Vanishing Theorem yields *all* such constraints.

10A The Final Constraint Set Lies in the Zero Level of the Energy-Momentum Map

Consider a Lagrangian field theory with gauge group $\mathcal{G}$. Recall from §6E that the final constraint set $\mathcal{C}_\tau$ consists of all initial data on the Cauchy surface Σ_τ which can be formally integrated to dynamical trajectories which satisfy Hamilton's equations. Throughout we suppose that the regularity assumption **A3** holds, viz. that $\mathcal{C}_\tau$ is a manifold and $\ker \omega_{\mathcal{C}_\tau}$ is a regular distribution. We will also need the following refinement of the notion of Lagrangian slicing:

A6 $\mathcal{G}$**-Slicings.** *For a classical field theory with gauge group $\mathcal{G}$ and configuration bundle Y, there exists a $\mathcal{G}$-slicing of Y.*

Our first main result is essentially a "Hamiltonian restatement" of the Vanishing Theorem:

Theorem 10A.1. *Suppose that the Euler–Lagrange equations are well-posed. Then*

$$\mathcal{C}_\tau \subset \mathcal{E}_\tau^{-1}(0) \tag{10A.1}$$

where $\mathcal{E}_\tau$ is the instantaneous energy-momentum map induced on $\mathcal{P}_\tau$.

Proof. Let $(\varphi, \pi) \in \mathcal{C}_\tau$. By **A4** and Theorem 6D.4(*ii*), there is a solution $\phi \in \mathcal{Y}$ of the Euler–Lagrange equations such that $\mathrm{can}_\tau(\phi) = (\varphi, \pi)$. Set $\sigma = \mathbb{F}\mathcal{L} \circ j^1\phi \circ \tau$ so that σ is a holonomic lift of (φ, π). Now apply the Vanishing Theorem 9A.2 to ϕ, obtaining for each $\xi \in \mathfrak{g}$,

$$\begin{aligned} 0 &= \int_\Sigma \tau^*(j^1\phi)^* J^{\mathcal{L}}(\xi) = \int_\Sigma \tau^*(j^1\phi)^* \mathbb{F}\mathcal{L}^* J(\xi) && \text{(by §4D)} \\ &= \int_\Sigma \sigma^* J(\xi) = \langle E_\tau(\sigma), \xi \rangle && \text{(by (7B.1))} \\ &= \langle \mathcal{E}_\tau(\varphi, \pi), \xi \rangle \end{aligned}$$

according to §7D, as σ is a holonomic lift of (φ, π). Thus $(\varphi, \pi) \in \mathcal{E}_\tau^{-1}(0)$. ■

This result shows that the conditions $\langle \mathcal{E}_\tau, \xi \rangle = 0$ are secondary initial value constraints. In fact:

Proposition 10A.2. *The components $\langle \mathcal{E}_\tau, \xi \rangle$ of the energy-momentum map are first class functions.*

Proof. We must show that

$$T\mathcal{C}_\tau^\perp \left[\langle \mathcal{E}_\tau, \xi \rangle \right] = 0 \tag{10A.2}$$

for each $\xi \in \mathfrak{g}$. We break the proof into three parts, depending upon whether or not ξ_X is tangent to Σ_τ, is transverse to Σ_τ, or is a combination.

First suppose that ξ_X is everywhere tangent to Σ_τ, so that $\xi \in \mathfrak{g}_\tau$. In this case, Proposition 8C.6 ff. asserts that $\xi_{\mathcal{C}_\tau} = \xi_{T^*\mathcal{Y}_\tau} | \, \mathcal{C}_\tau$ satisfies

$$\xi_{\mathcal{C}_\tau} \,\lrcorner\, \omega_\tau = \mathbf{d}\langle \mathcal{E}_\tau, \xi \rangle \tag{10A.3}$$

along $\mathcal{C}_\tau$. Furthermore, we know that $\xi_{\mathcal{C}_\tau}$ is tangent to $\mathcal{C}_\tau$ so that (10A.3), when evaluated on vectors in $T\mathcal{C}_\tau^\perp$, yields (10A.2).

Next suppose that ξ_X is everywhere transverse to the surface Σ_τ. Then Corollary 7D.2(i) states that $\langle \mathcal{E}_\tau, \xi \rangle = -H_{\tau,\xi}$. Then (10A.2) is satisfied by virtue of Corollary 6E.3.

Finally, consider the intermediate case when ξ_X is neither everywhere tangent nor everywhere transverse to Σ_τ. Utilizing assumption **A6**, let ζ be the generating vector field of any $\mathcal{G}$-slicing of Y; then $\zeta \in \mathfrak{g}$ and $\zeta_X \pitchfork \Sigma_\tau$. Along Σ_τ, we may decompose

$$\xi_X = \xi_X^\parallel + f\zeta_X$$

where $\xi_X^\parallel$ is tangent to Σ_τ and f is some function on Σ_τ. Choose any $k > \max_{x\in\Sigma_\tau} |f(x)|$ (such a number exists as Σ_τ is compact), and define $\vartheta = k\zeta - \xi$. Then $\vartheta \in \mathfrak{g}$ and $\vartheta \pitchfork \Sigma_\tau$ so we may write $\xi = k\zeta - \vartheta$ as the difference of two transverse gauge generators. Since $\langle \mathcal{E}_\tau, \xi \rangle = k\langle \mathcal{E}_\tau, \zeta \rangle - \langle \mathcal{E}_\tau, \vartheta \rangle$, the desired result follows from the transverse case proved above and the comment immediately following Corollary 6E.12. ■

In summary: *The components $\langle \mathcal{E}_\tau, \xi \rangle$ of the energy-momentum map are first class secondary initial value constraints.*

Remark 10A.3. Recall that for secondary constraints we measure class relative to $(\mathcal{P}_\tau, \omega_\tau)$, cf. §6E. ♦

Remark 10A.4. This result combined with Proposition 6E.1gives Proposition 8C.1. Theorem 8C.2 now follows either from Propositions 8C.1 and 6E.8(i) or from Propositions 10A.2 and 6E.8(iii). ♦

Remark 10A.5. Theorem 10A.1 together with Corollary 7D.2(i) show that each Hamiltonian $H_{\tau,\xi}$ vanishes identically on $\mathcal{C}_\tau$. Thus on $\mathcal{C}_\tau$ the Hamilton equations (6E.11)pull back to $\mathbf{i}_X\omega_{\mathcal{C}_\tau} = 0$, that is to say, $X \in \mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$; in this sense the evolution is "purely gauge." This is one of the hallmarks of parametrized theories (in which all fields are variational), cf. Kuchař [1973]. ♦

Remark 10A.6. When working with background theories, Theorem 10A.1 remains valid as long as **A6** is satisfied. So does Proposition 10A.2 (but now the tangential case is the only relevant one). Proposition 10A.2 also holds in theories in which there are non-variational fields, but in this context Theorem 10A.1 fails (because the first Noether theorem 4D.2 does) and so there is no reason why, for instance, the Hamiltonian should vanish on $\mathcal{C}_\tau$. See Example **b** at the end of this chapter for more details. ♦

With a slight shift in viewpoint, it is possible to substantially weaken the well-posedness hypothesis in Theorem 10A.1. Consistent with our general philosophy, suppose we know *ab initio* that $\mathcal{G}$ "acts" on $\mathcal{C}_\tau$ by gauge transformations in the sense that Theorem 8C.2 holds (in particular, we might assume that the gauge group is full). Then for each $\xi \in \mathfrak{g}$, the gauge vector field $\xi_{\mathcal{C}_\tau} \in \mathfrak{X}(\mathcal{C}_\tau) \cap \mathfrak{X}(\mathcal{C}_\tau)^\perp$ satisfies equation (8C.1):

$$(\xi_{\mathcal{C}_\tau} \lrcorner\, \omega_\tau - \mathbf{d}\langle \mathcal{E}_\tau, \xi\rangle) \,|\, \mathcal{C}_\tau = 0.$$

Evaluate this equation on $T\mathcal{C}_\tau$; since pointwise $\xi_{\mathcal{C}_\tau} \in T\mathcal{C}_\tau \cap T\mathcal{C}_\tau^\perp$, we find that

$$T\mathcal{C}_\tau[\langle \mathcal{E}_\tau, \xi\rangle] = 0.$$

Thus each component $\langle \mathcal{E}_\tau, \xi\rangle$ is constant along $\mathcal{C}_\tau$. Now suppose that at least one initial data point $(\varphi, \pi) \in \mathcal{C}_\tau$ can be integrated to a spacetime solution ϕ of the Euler–Lagrange equations. Applying the Vanishing Theorem to ϕ, it follows that each $\langle \mathcal{E}_\tau(\varphi, \pi), \xi\rangle$ is zero on $\mathcal{C}_\tau$. We have therefore proven the following variant of Theorem 10A.1.

Theorem 10A.7. *Let $\mathcal{G}$ be the gauge group for the Lagrangian field theory under consideration. Suppose that the conclusion of Theorem 8C.2 holds and that at least one initial data point in $\mathcal{C}_\tau$ admits a finite time evolution. If $\mathcal{C}_\tau$ is connected, then $\mathcal{C}_\tau \subset \mathcal{E}_\tau^{-1}(0)$.*

Thus we obtain the desired inclusion with the assumption that only *one*, as opposed to *every*, initial data set is integrable. This amounts to requiring that the field theory be non-trivial, i.e., Sol $\neq \varnothing$.

Remark 10A.8. If $\mathcal{C}_\tau$ is not connected, then Theorem 10A.7 applies only to those components which contain well-posed initial data. Other components can be ignored. ♦

10B The Energy-Momentum Theorem

We have shown that the components of the instantaneous energy-momentum map are first class secondary constraints. Our goal now is to show that $\mathcal{E}_\tau$ contains *all* such constraints when the gauge group is full.

To begin, we prove equality in either of Theorems 10A.1 or 10A.7 under the hypothesis that the field theory is first class. Throughout this section we make whatever assumptions are necessary for Theorem 10A.1 (or its variant

Theorem 10A.7) to hold. We also suppose that $\mathcal{E}_\tau^{-1}(0)$ is a submanifold of $\mathcal{P}_\tau$. (Remark 10B.8 below discusses the singular case.)

Theorem 10B.1. (Energy-Momentum Theorem, Version I) *If $\mathcal{G}$ is full, all secondary constraints are first class, and $\mathcal{E}_\tau^{-1}(0)$ is connected, then*

$$\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0). \tag{10B.1}$$

Proof. To say that all secondary constraints are first class is the same as saying that $\mathcal{C}_\tau$ is coisotropic in $\mathcal{P}_\tau$: $T\mathcal{C}_\tau^\perp \subset T\mathcal{C}_\tau$. Then at every point of $\mathcal{C}_\tau$ **A5** requires

$$\mathfrak{g}_{\mathcal{C}_\tau} = T\mathcal{C}_\tau^\perp. \tag{10B.2}$$

Next, suppose $V \in \ker T\mathcal{E}_\tau$ at a point of $\mathcal{C}_\tau$. Then equation (8C.1) yields $\omega_\tau(\xi_{\mathcal{C}_\tau}, V) = 0$ for all $\xi \in \mathfrak{g}$. This means that

$$\ker T\mathcal{E}_\tau \,|\, \mathcal{C}_\tau \subset (\mathfrak{g}_{\mathcal{C}_\tau})^\perp = T\mathcal{C}_\tau \subset T\mathcal{E}_\tau^{-1}(0) \,|\, \mathcal{C}_\tau \tag{10B.3}$$

by (10B.2), Lemma 6E.7, and (10A.1). Since obviously $T\mathcal{E}_\tau^{-1}(0) \subset \ker T\mathcal{E}_\tau$, we conclude that

$$\ker T\mathcal{E}_\tau = T\mathcal{E}_\tau^{-1}(0)$$

along $\mathcal{C}_\tau$. Substituting into (10B.3) gives finally

$$T\mathcal{C}_\tau = T\mathcal{E}_\tau^{-1}(0) \,|\, \mathcal{C}_\tau.$$

The inverse function theorem shows that $\mathcal{C}_\tau$ is open in $\mathcal{E}_\tau^{-1}(0)$. But $\mathcal{C}_\tau$ is closed by its very construction (being given by the vanishing of constraints), so (10B.1) follows from the connectedness of $\mathcal{E}_\tau^{-1}(0)$. ■

Remark 10B.2. In the event that $\mathcal{E}_\tau^{-1}(0)$ is not connected, $\mathcal{C}_\tau$ is a union of components of $\mathcal{E}_\tau^{-1}(0)$. In practice, the connectedness hypothesis is not a problem; Example **a** in §11D provides an illustration. ♦

Remark 10B.3. An examination of the first part of the proof shows that if $\mathcal{G}$ is full and $\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0)$ is a coisotropic submanifold of $\mathcal{P}_\tau$, then zero must be a ***weakly regular value*** of $\mathcal{E}_\tau$; that is, in addition to $\mathcal{E}_\tau^{-1}(0)$ being smooth,

$$\ker T\mathcal{E}_\tau = T\mathcal{E}_\tau^{-1}(0). \tag{10B.4}$$

However, the operators $\mathbf{D}\mathcal{E}_\tau$ are generally elliptic in the appropriate sense, so that one can sometimes prove, relative a suitably chosen Sobolev topology, that

0 is a regular value. See Arms et al. [1982] and Interlude IV for more information and references on this point. We usually shall not pursue the distinction in this work for simplicity. ♦

Here is another way to obtain equality in (10A.1) which assumes that $\mathcal{E}_\tau^{-1}(0)$, rather than $\mathcal{C}_\tau$, is coisotropic.

Theorem 10B.4. (Energy-Momentum Theorem, Version II) *If* $\mathcal{E}_\tau^{-1}(0)$ *is coisotropic, then* $\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0)$.

Proof. Let $\zeta \in \mathfrak{g}$ with $\zeta_X \pitchfork \Sigma_\tau$; such a ζ exists by virtue of **A6**. By Corollary 7D.2(i), the corresponding Hamiltonian $-H_{\tau,\zeta}$ is a component of $\mathcal{E}_\tau$, so

$$T\mathcal{E}_\tau^{-1}(0)[H_{\tau,\zeta}] = 0.$$

As $\mathcal{E}_\tau^{-1}(0)$ is coisotropic, this implies that

$$T\mathcal{E}_\tau^{-1}(0)^\perp[H_{\tau,\zeta}] = 0.$$

The desired result now follows from Corollary 6E.3,Corollary 6E.12,(along with the comment afterwards), and Theorem 10A.1. ■

Theorem 10B.1 is perhaps more theoretically appealing than Theorem 10B.4, but the latter has three important advantages over the former. First, it requires no *a priori* knowledge of $\mathcal{C}_\tau$. Second, its hypotheses are straightforward (although not necessarily trivial!) to verify in practice. Last, it eliminates the necessity of having to worry about the connectedness of $\mathcal{E}_\tau^{-1}(0)$. Later in §13C we will see that one can check for equality of $\mathcal{C}_\tau$ and $\mathcal{E}_\tau^{-1}(0)$ by equation counting.

Using Proposition 10A.2, we obtain a partial converse to Theorem 10B.4.

Proposition 10B.5. *If zero is a weakly regular value of* $\mathcal{E}_\tau$ *and* $\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0)$, *then* $\mathcal{E}_\tau^{-1}(0)$ *is coisotropic.*

Proof. By hypothesis and equation (10A.2),

$$T\mathcal{E}_\tau^{-1}(0)^\perp\big[\langle \mathcal{E}_\tau, \xi\rangle\big] = 0$$

for all $\xi \in \mathfrak{g}$. In other words,

$$T\mathcal{E}_\tau^{-1}(0)^\perp \subset \ker T\mathcal{E}_\tau = T\mathcal{E}_\tau^{-1}(0)$$

by the weak regularity of 0. ■

In Theorem 10B.1 we had to assume that $\mathcal{G}$ was full, but not in Theorem 10B.4. This is because of the next result which, in view of Remark 10B.3 above, can be considered a converse to Theorem 10B.1.

Corollary 10B.6. *If zero is a weakly regular value of $\mathcal{E}_\tau$ and $\mathcal{E}_\tau^{-1}(0)$ is coisotropic, then $\mathcal{G}$ is full.*

Proof. By Theorem 10B.4, $\mathcal{E}_\tau^{-1}(0) = \mathcal{C}_\tau$. So we must show pointwise that

$$\mathfrak{g}_{\mathcal{E}_\tau^{-1}(0)} = T\mathcal{E}_\tau^{-1}(0)^\perp.$$

From (8C.1) and weak regularity

$$\mathfrak{g}^\perp_{\mathcal{E}_\tau^{-1}(0)} = \ker T\mathcal{E}_\tau = T\mathcal{E}_\tau^{-1}(0).$$

Taking polars we obtain

$$\mathfrak{g}^{\perp\perp}_{\mathcal{E}_\tau^{-1}(0)} = T\mathcal{E}_\tau^{-1}(0)^\perp. \tag{10B.5}$$

Thus we will be done if we can show that $\mathfrak{g}^{\perp\perp}_{\mathcal{E}_\tau^{-1}(0)} = \mathfrak{g}_{\mathcal{E}_\tau^{-1}(0)}$.

To this end Lemma 6E.7 gives

$$\mathfrak{g}^{\perp\perp}_{\mathcal{E}_\tau^{-1}(0)} = \mathfrak{g}_{\mathcal{E}_\tau^{-1}(0)} + T\mathcal{P}_\tau^\perp \tag{10B.6}$$

along $\mathcal{E}_\tau^{-1}(0)$. Now by (10B.5) and the assumption that $\mathcal{E}_\tau^{-1}(0)$ is coisotropic, $\mathfrak{g}^{\perp\perp}_{\mathcal{E}_\tau^{-1}(0)}$ is tangent to $\mathcal{E}_\tau^{-1}(0)$, and by definition so is $\mathfrak{g}_{\mathcal{E}_\tau^{-1}(0)}$. From (10B.6) we conclude that $T\mathcal{P}_\tau^\perp \subset T\mathcal{E}_\tau^{-1}(0)$. But then $T\mathcal{P}_\tau^\perp = T\mathcal{P}_\tau^\perp \cap T\mathcal{E}_\tau^{-1}(0) \subset \mathfrak{g}_{\mathcal{E}_\tau^{-1}(0)}$ (cf. Remark 8C.3) and so (10B.6) shows that $\mathfrak{g}^{\perp\perp}_{\mathcal{E}_\tau^{-1}(0)} = \mathfrak{g}_{\mathcal{E}_\tau^{-1}(0)}$. ■

These last two results show that the fullness of $\mathcal{G}$ is a necessary condition for equality in (10A.1). It is not sufficient, however, since clearly neither (10B.1) nor the converse of Corollary 10B.6 can hold in the presence of second class secondary constraints. Notice also that Corollary 10B.6 can be used to *show* that $\mathcal{G}$ is full.

Thus, roughly speaking, we have:

$$\mathcal{G} \text{ full } + \ \mathcal{C}_\tau \text{ coisotropic } \Leftrightarrow\ \mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0) \ \Leftrightarrow\ \mathcal{E}_\tau^{-1}(0) \text{ coisotopic.}$$

Remark 10B.7. With regard to the hypotheses of Theorem 10B.4 and Corollary 10B.6, observe that $\mathcal{E}_\tau^{-1}(0)$ is *not* necessarily coisotropic. This is because $\mathcal{E}_\tau$ is not a true momentum map, but rather an energy-momentum map. (Of course, the zero levels of momentum maps for cotangent actions—and for arbitrary actions, provided zero is a weakly regular value—are coisotropic; see Arms et al. [1989] and Abraham and Marsden [1978]) We illustrate this point in our discussion of Palatini gravity in Part V.

Note that Proposition 10A.2, which states that the components of $\mathcal{E}_\tau$ are first class constraints, does *not* imply that $\mathcal{E}_\tau^{-1}(0)$ is globally coisotropic. Rather, it only shows that $\mathcal{E}_\tau^{-1}(0)$ is coisotropic *along* $\mathcal{C}_\tau$. ♦

Remark 10B.8. Throughout this chapter we have been treating $\mathcal{C}_\tau$ and $\mathcal{E}_\tau^{-1}(0)$ as if they were smooth manifolds. In practice, of course, they are not, except in special (linear) cases such as electromagnetism or abelian Chern–Simons theory (cf. Remark 6E.4). However, we emphasize that Theorem 10A.1 holds regardless, as does Theorem 10B.1 provided certain reasonable conditions are met. Suppose both $\mathcal{E}_\tau^{-1}(0)$ and $\mathcal{C}_\tau$ are the closures of their smooth points, denoted $S(\mathcal{E}_\tau^{-1}(0))$ and $S(\mathcal{C}_\tau)$, and that

$$S(\mathcal{C}_\tau) = S(\mathcal{E}_\tau^{-1}(0)) \cap \mathcal{C}_\tau; \tag{10B.7}$$

in other words, the smooth points of $\mathcal{C}_\tau$ are also smooth points of $\mathcal{E}_\tau^{-1}(0)$. If the set $S(\mathcal{E}_\tau^{-1}(0))$ is connected, then the argument in the proof of Theorem 10B.1 shows that $S(\mathcal{C}_\tau)$ is open in $S(\mathcal{E}_\tau^{-1}(0))$. But (10B.7) implies that it is closed in addition, so $S(\mathcal{C}_\tau) = S(\mathcal{E}_\tau^{-1}(0))$, and thus we recover (10B.1) upon taking closures. In a similar manner, one can show that Theorem 10B.4 etc. hold, using the groundwork laid in Arms et al. [1990]. ♦

Remark 10B.9. *Given* that $\mathcal{G}$ is the gauge group of a particular classical field theory, then in view of Theorem 10B.1 the fullness assumption **A5** provides a sufficient condition for the validity of the "Dirac conjecture" that *all* first class secondary constraints generate gauge transformations. Gotay [1983] contains more information. ♦

10C First Class Secondary Constraints

Either version of the Energy-Momentum Theorem requires that the field theory under consideration be first class in an appropriate sense. Here we consider the general case.

To prepare for this, we now give an algebraic restatement of our results. Let $I(\mathcal{C}_\tau)$ be the ideal[27] in $\mathcal{F}(\mathcal{P}_\tau)$ consisting of secondary constraints (i.e., smooth functions on $\mathcal{P}_\tau$ vanishing on $\mathcal{C}_\tau$), and denote by $(\mathcal{E}_\tau)$ the ideal generated by

[27] All ideals are multiplicative.

the components $\langle \mathcal{E}_\tau, \xi \rangle$ of $\mathcal{E}_\tau$ as ξ ranges over $\mathfrak{g}$. From either Theorem 10A.1 or Theorem 10A.7 we have that $(\mathcal{E}_\tau) \subset I(\mathcal{C}_\tau)$.

Proposition 10C.1. *Suppose (10B.1) holds and that zero is a regular value of $\mathcal{E}_\tau$, so that all secondary constraints are first class. Then*

$$I(\mathcal{C}_\tau) = (\mathcal{E}_\tau). \tag{10C.1}$$

In other words, if f is a secondary constraint, then

$$f = \sum_\alpha g_\alpha \langle \mathcal{E}_\tau, \xi_\alpha \rangle \tag{10C.2}$$

for some smooth functions g_α, where $\{\xi_\alpha\}$ forms a basis for $\mathfrak{g}$ in a topology appropriate to the analysis needed for the particular application (usually this means a Sobolev topology, so (10C.2) will generally be a countably infinite sum).

Proof. By (10B.1), it suffices to demonstrate that $f \in (\mathcal{E}_\tau)$ iff $f \,|\, \mathcal{E}_\tau^{-1}(0) = 0$. The obverse is immediate. For the converse, suppose $f \mid \mathcal{E}_\tau^{-1}(0) = 0$. We will prove that there exists open sets U about each $(\varphi, \pi) \in \mathcal{P}_\tau$, such that $f \,|\, U \in (\mathcal{E}_\tau) \,|\, U$. The Proposition then follows by patching these local results together with a partition of unity.[28]

There are two cases to consider: $(\varphi, \pi) \notin \mathcal{E}_\tau^{-1}(0)$ and $(\varphi, \pi) \in \mathcal{E}_\tau^{-1}(0)$. Fix a basis $\{\xi_\alpha\}$ for $\mathfrak{g}$. For the first case, choose U such that $U \cap \mathcal{E}_\tau^{-1}(0) = \varnothing$. By shrinking U if necessary, we can find an $\hat{\alpha}$ such that $\langle \mathcal{E}_\tau, \xi_{\hat{\alpha}} \rangle \neq 0$ on U. Then $g = (f \,|\, U)/\langle \mathcal{E}_\tau, \xi_{\hat{\alpha}} \rangle$ is smooth on U, and

$$f \,|\, U = g \langle \mathcal{E}_\tau, \xi_{\hat{\alpha}} \rangle \in (\mathcal{E}_\tau) \,|\, U.$$

Secondly, let $(\varphi, \pi) \in \mathcal{E}_\tau^{-1}(0)$. Since 0 is a regular value of $\mathcal{E}_\tau$, a neighborhood U of (φ, π) is isomorphic to a neighborhood V of the origin in $\ker T\mathcal{E}_\tau \times \mathfrak{g}^* = T\mathcal{E}_\tau^{-1}(0) \times \mathfrak{g}^*$ in such a way that $\mathcal{E}_\tau(x, \mu) = \mu$ on V (cf. Theorem 2.5.15 in Abraham et al. [1988]). By Taylor's theorem with remainder, $f \mid \big(V \cap (T\mathcal{E}_\tau^{-1}(0) \times \{0\})\big) = 0$ implies that $f(x, \mu) = \sum_\alpha g_\alpha \mu^\alpha$ on V for some smooth functions g_α, where $\{\mu^\alpha\}$ is the dual basis to $\{\xi_\alpha\}$. Thus on U, we have $f \,|\, (U \cap \mathcal{E}_\tau^{-1}(0)) = 0$ implies that $f \,|\, U = \sum_\alpha g_\alpha \langle \mathcal{E}_\tau, \xi_\alpha \rangle$. ■

[28] If partitions of unity are not available, then one must localize (10C.1) to sufficiently small open sets in $\mathcal{P}_\tau$.

When second class secondary constraints appear, neither (10B.1) nor (10C.1) will remain valid. But we can recover their essential content by utilizing the well-known "Dirac bracket construction" to eliminate these constraints.

Here is how this works in the case at hand. Apply the standard Dirac bracket construction as developed in Śniatycki [1974] to $\mathcal{C}_\tau$ viewed as a submanifold of $T^*\mathcal{Y}_\tau$, thereby obtaining a symplectic submanifold $\mathcal{V}_\tau \subset T^*\mathcal{Y}_\tau$ containing $\mathcal{C}_\tau$ as a coisotropic submanifold. Set $\mathcal{W}_\tau = \mathcal{V}_\tau \cap \mathcal{P}_\tau$; then it is readily verified that (i) $\mathcal{W}_\tau$ is a second class submanifold of $\mathcal{C}_\tau$, and (ii) $\mathcal{C}_\tau$ is also a coisotropic submanifold of $\mathcal{W}_\tau$. The point is that, viewing such a choice of a ***Dirac manifold*** $\mathcal{W}_\tau$ as the ambient space and $\mathcal{C}_\tau$ as the final constraint set *therein*, all the secondary constraints are now purely first class.

Setting $\mathcal{K}_\tau = \mathcal{E}_\tau \,|\, \mathcal{W}_\tau$, we may therefore apply Theorem 10B.1 and Proposition 10C.1 verbatim to $\mathcal{C}_\tau \subset \mathcal{W}_\tau$ and $\mathcal{K}_\tau$, with the result that

$$I(\mathcal{C}_\tau) = (\mathcal{K}_\tau), \tag{10C.3}$$

the ideals now being taken in $\mathcal{F}(\mathcal{W}_\tau)$.

Let $I^F(\mathcal{C}_\tau)$ denote the ideal in $\mathcal{F}(\mathcal{P}_\tau)$ consisting of first class constraints, and let $I(\mathcal{W}_\tau)$ be the ideal of all smooth functions on $\mathcal{P}_\tau$ vanishing on $\mathcal{W}_\tau$ (so that $I(\mathcal{W}_\tau)$ is "generated by the second class constraints"). Set $I^F(\mathcal{W}_\tau) = I^F(\mathcal{C}_\tau) \cap I(\mathcal{W}_\tau)$. Pulling (10C.3) back to $\mathcal{P}_\tau$, we finally have:

Theorem 10C.2. *Suppose that zero is a regular value of $\mathcal{E}_\tau$, and that either (i) $\mathcal{G}$ is full and $\mathcal{E}_\tau^{-1}(0) \cap \mathcal{W}_\tau$ is connected, or (ii) $\mathcal{E}_\tau^{-1}(0) \cap \mathcal{W}_\tau$ is coisotropic in $\mathcal{W}_\tau$. Then*

$$I^F(\mathcal{C}_\tau) \equiv (\mathcal{E}_\tau) \mod I^F(\mathcal{W}_\tau).$$

Proof. Let $f \in I^F(\mathcal{C}_\tau)$. Then $f \,|\, \mathcal{W}_\tau$ vanishes on $\mathcal{C}_\tau$, so our assumptions imply (10C.3), which gives

$$f \,|\, \mathcal{W}_\tau = \sum_\alpha \tilde{g}_\alpha \langle \mathcal{K}_\tau, \xi_\alpha \rangle$$

for some smooth functions $\tilde{g}_\alpha$ on $\mathcal{W}_\tau$. We may therefore write

$$f = \sum_\alpha g_\alpha \langle \mathcal{E}_\tau, \xi_\alpha \rangle + \psi$$

on $\mathcal{P}_\tau$, where the g_α extend the $\tilde{g}_\alpha$ and $\psi \,|\, \mathcal{W}_\tau = 0$ with ψ first class. ■

Theorem 10C.2(i) is optimal insofar as it guarantees—modulo the connectedness of $\mathcal{E}_\tau^{-1}(0) \cap \mathcal{W}_\tau$—that one can recover all the "important" first class secondary constraints from $\mathcal{E}_\tau$ without having to determine beforehand the class of all the constraints. The latter, of course, would necessitate running through the entire constraint algorithm. This works provided the gauge group is full; as part of our general philosophy (cf. the discussion in §8C), we assume that this is known at the outset, as is the case in all our examples.

The only constraints that cannot be obtained in this fashion are those belonging to $I^F(\mathcal{W}_\tau) = I^F(\mathcal{C}_\tau) \cap I(\mathcal{W}_\tau)$; since these first class secondaries are generated by the second class secondaries, they are clearly irrelevant insofar as constraint theory is concerned. In fact, such constraints ψ are ***ineffective*** in the sense that $\mathbf{d}\psi \,|\, \mathcal{C}_\tau = 0$.

To see this, let $\psi \in I^F(\mathcal{C}_\tau) \cap I(\mathcal{W}_\tau)$, and let $U \subset \mathcal{P}_\tau$ be an open set such that $U \cap \mathcal{C}_\tau \neq \varnothing$. Let $\{\chi_\alpha\}$ be a functionally independent set of second class constraints defining $U \cap \mathcal{W}_\tau$ in U. By our assumptions, we may write

$$\psi \,|\, U = \sum_\alpha h_\alpha \chi_\alpha \tag{10C.4}$$

for some smooth functions h_α. As the χ_α are second class, for each α there is a vector field $\mathcal{V}_\alpha \in \mathfrak{X}(U \cap \mathcal{C}_\tau)^\perp$ such that $\mathcal{V}_\alpha[\chi_\beta] = \delta_{\alpha\beta}$ on some shrunken neighborhood $V \subset U$. On the other hand, ψ is first class, whence $\mathcal{V}_\alpha[\psi] = 0$ on $V \cap \mathcal{C}_\tau$. Taking differentials in (10C.4), restricting to V, and evaluating on each $\mathcal{V}_\alpha$ in turn, we see that $h_\alpha \,|\, (V \cap \mathcal{C}_\tau) = 0$, whence the desired result.

Of course, when there are no second class secondaries, Theorem 10C.2 reduces to Proposition 10C.1 in which case we get all the secondaries via $\mathcal{E}_\tau$.

Remark 10C.3. The global geometry and topology of the Dirac manifold $\mathcal{W}_\tau$ are not uniquely determined by our constructions. However, our results are insensitive to the global structure of $\mathcal{W}_\tau$. ♦

Remark 10C.4. In general, $\mathcal{W}_\tau$ will not be symplectic (in contrast to $\mathcal{V}_\tau$), since the surrounding space $\mathcal{P}_\tau$ itself need not be symplectic. (But $\mathcal{W}_\tau$ is always second class.) Even so, $I^F(\mathcal{W}_\tau)$ is usually nonzero. The reason is that if f is a second class constraint, then f^2 is a first class constraint. ♦

Remark 10C.5. It may happen that $\mathcal{C}_\tau$ is symplectic. In this case our results still apply—albeit trivially—with $\mathcal{G}$ the zero group, so $\mathcal{E}_\tau = 0$, $\mathcal{W}_\tau = \mathcal{C}_\tau$, and Theorem 10C.2 reduces to a tautology. ♦

Remark 10C.6. If $\mathcal{C}_\tau$ is not a smooth manifold, we localize our results to $\mathcal{F}(\mathcal{U})$, for open sets $\mathcal{U}$ in $\mathcal{P}_\tau$ such that $\mathcal{U} \cap \mathcal{C}_\tau \subset S(\mathcal{C}_\tau)$. ♦

The Dirac bracket construction enables us to work on $\mathcal{W}_\tau$ instead of $\mathcal{P}_\tau$. This is possible by the following result, which encapsulates a key property of Dirac manifolds.

Proposition 10C.7. *Suppose* $\mathcal{C} \subset (\mathcal{W}, \varpi) \subset (\mathcal{P}, \omega)$, *with the second inclusion being presymplectic. Let* H *be a Hamiltonian on* $\mathcal{P}$. *If there exists a vector field* $X \in \mathfrak{X}(\mathcal{C})$ *such that*

$$(\mathbf{i}_X \omega - \mathbf{d}H) \,|\, \mathcal{C} = 0 \tag{10C.5}$$

then

$$(\mathbf{i}_X \varpi - \mathbf{d}h) \,|\, \mathcal{C} = 0, \tag{10C.6}$$

where h *is the pull-back of* H *to* $\mathcal{W}$. *Conversely, assume that* $\mathcal{W}$ *is a second class submanifold of* $\mathcal{P}$ *and let* h *be a Hamiltonian on* $\mathcal{W}$. *If there exists a vector field* $X \in \mathfrak{X}(\mathcal{C})$ *such that* (10C.6) *holds, then* X *satisfies* (10C.5) *for some extension* H *of* h *to* $\mathcal{P}$.

Proof. Since X is assumed to be tangent to $\mathcal{C}$, (10C.5) $\Rightarrow$ (10C.6) by pulling the former back to $\mathcal{W}$.

The opposite implication is subtler. Let $\mathcal{U} \subset \mathcal{P}$ be an open set with $\mathcal{U} \cap \mathcal{C} \neq \varnothing$, chosen so that there exists a basis of constraints $\{\chi_\alpha\}$ defining $\mathcal{U} \cap \mathcal{W} \subset \mathcal{U}$. By our assumptions and Proposition 6E.8(vi) the χ_α are second class. Consequently, for each α we may find (shrinking $\mathcal{U}$ as necessary) a collection of vector fields $\{V_\alpha\}$ on $\mathcal{U}$ which forms a basis for $\mathfrak{X}(\mathcal{U} \cap \mathcal{W})^\perp$ such that $V_\alpha[\chi_\beta] = \delta_{\alpha\beta}$. By again shrinking $\mathcal{U}$, if required, we may suppose there exist extensions $\tilde{h}$ of h and $\tilde{X}$ of X satisfying (10C.6) to $\mathcal{U}$.

Now if H is any extension of h to $\mathcal{U}$, we may write $H = \tilde{h} + \sum_\alpha g_\alpha \chi_\alpha$ for some smooth functions g_α. We show how to choose H in such a way that (10C.5) holds along $\mathcal{U} \cap \mathcal{C}$.

First observe that (10C.6) shows that (10C.5) holds when evaluated on $T\mathcal{W}$ along $\mathcal{U} \cap \mathcal{C}$, regardless of the choice of H. Next, evaluate (10C.5) on V_α. Then $\omega(\tilde{X}, V_\alpha) \,|\, (\mathcal{U} \cap \mathcal{C}) = 0$, and we may arrange $dH(V_\alpha) \,|\, (\mathcal{U} \cap \mathcal{C}) = 0$ by fixing $g_\alpha = -V_\alpha[\tilde{h}]$. With this choice, (10C.5) also holds when evaluated on $T\mathcal{W}^\perp$ along $\mathcal{U} \cap \mathcal{C}$. But then the second class condition $T\mathcal{W} + T\mathcal{W}^\perp = T_\mathcal{W}\mathcal{P}$ establishes (10C.5). It only remains to patch these locally defined extensions together with an appropriate partition of unity. ∎

Thus Hamilton's equations on the primary constraint set have exactly the same solutions as on the Dirac manifold: the two arenas for dynamics are entirely equivalent. This Dirac bracket construction can save substantial computational effort, and we shall repeatedly use this construction in the remainder of the book.

Examples

a Particle Mechanics. In the case of the standard relativistic free particle, $\mathcal{E}_t = 0$ and there are no secondary constraints. However, there is a formulation of relativistic dynamics in which the free particle does have secondary constraints; we will discuss this in Chapter 11.

b Electromagnetism. Most of the theory developed in this section applies only to parametrized field theories in which all fields are variational; nonetheless, the conclusions all hold for background electromagnetism (cf. Remark 10A.6). Indeed, we know from §8C that $\mathcal{F}(X)$ is full and from §6E that the divergence constraint (6E.21) is first class. Since clearly

$$\mathcal{E}_\tau^{-1}(0) = \{(A, \mathfrak{E}) \,|\, D_i \mathfrak{E}^i = 0\}$$

is connected, we conclude from Theorem 10B.1 that $\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0)$. Since furthermore the divergence constraint is linear, zero is a regular value of $\mathcal{E}_\tau$, and Proposition 10C.1 then shows that $I(\mathcal{C}_\tau) = (D_i \mathfrak{E}^i)$.

On the other hand, when the metric is treated parametrically (but is still nonvariational), our results are no longer valid. In particular, (10A.1) fails; in fact, we see from (7D.12) that now

$$\tilde{\mathcal{E}}_\tau^{-1}(0) \subset \tilde{\mathcal{C}}_\tau$$

strictly. The reversal of the inclusion here is due to $\tilde{\mathcal{G}}$ being overfull, as explained in Example 8C.**b**. Of course, in this case $\tilde{\mathcal{E}}_\tau^{-1}(0)$ is not coisotropic.

c A Topological Field Theory. From Examples **c** of §6E and §9B, we see that $\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0)$. This could have been anticipated on the basis of either of Theorems 10B.1 or 10B.4; note, for instance, that $\mathcal{E}_\tau^{-1}(0)$ given by $D_{[1} A_{2]} = 0$ is coisotropic. As this spatial flatness constraint is linear, Proposition 10C.1 may be applied with the result that $I(\mathcal{C}_\tau) = (D_{[1}A_{2]})$.

If we took the gauge group to be just $\mathrm{Diff}(X)$, which is not full, then from the computations in Example 9B.**c** we would have $\mathcal{E}_\tau = 0$, whence $\mathcal{C}_\tau \subset \mathcal{E}_\tau^{-1}(0) = \mathcal{P}_\tau$ strictly. Thus there are no supermomentum or superhamiltonian-type constraints for Chern–Simons theory, even though it has the entire spacetime diffeomporhism group as part of its gauge group.

d Bosonic Strings. The instantaneous energy-momentum map is given by (7D.17). Thus

$$\mathcal{E}_\tau^{-1}(0) = \big\{(\varphi, h, \pi) \mid \pi^2 + D\varphi^2 = 0,\ \pi \cdot D\varphi = 0\big\}$$

coincides with the final constraint set computed in Example **d** of §6E, as expected.

Interestingly, (10B.1) would remain valid if instead for $\mathcal{G}$ we used the (full) subgroup $\mathrm{Diff}(X)$ of $\mathcal{G}$. The difference between these two gauge groups—not apparent on the level of secondary constraints—will appear in the next chapter.

The bosonic string example is the first one in which the distinction between the energy-momentum map and the momentum map becomes crucial (insofar as identifying secondary constraints is concerned). While $\mathcal{J}_\tau$ picks up the supermomentum constraint $\pi \cdot D\varphi = 0$, it misses the superhamiltonian constraint $\pi^2 + D\varphi^2 = 0$. This is because the latter constraint is associated with that part of the spacetime diffeomorphism group which moves Σ_τ in X, and only $\mathcal{E}_\tau$ is sensitive to $\mathcal{G}/\mathcal{G}_\tau$. The momentum map also misses the superhamiltonian constraints (11D.10) and (14B.42) of the Polyakov particle and Palatini gravity, respectively.

11 Primary Constraints and the Momentum Map

The results of Chapter 10 enabled us to recover first class *secondary* constraints from the energy-momentum map by means of the Vanishing Theorem. Here we show that first class *primary* constraints can be obtained in a similar manner, by equating to zero the "momentum map" for a certain foliation $\dot{\mathcal{G}}_\tau$ derived from the gauge group $\mathcal{G}$. We begin with a description of $\dot{\mathcal{G}}_\tau$ and its "action."

11A Motivation

We know that the vanishing of the energy-momentum map $\mathcal{E}_\tau : \mathcal{P}_\tau \to \mathfrak{g}^*$ yields first class secondary constraints; moreover, for first class systems, $\mathcal{C}_\tau = \mathcal{E}_\tau^{-1}(0)$. It is natural to ask whether the first class primary constraints can be recovered in an analogous fashion. The obvious counterpart of $\mathcal{E}_\tau$ in this context is the momentum map $\mathcal{J}_\tau : T^*\mathcal{Y}_\tau \to \mathfrak{g}_\tau^*$ for the action of $\mathcal{G}_\tau$ on $T^*\mathcal{Y}_\tau$. (Recall from §7B that $\mathcal{G}_\tau$ consists of those elements of the gauge group $\mathcal{G}$ that stabilize the hypersurface Σ_τ and that $\mathfrak{g}_\tau$ is its Lie algebra.) Now, in coordinates adapted to Σ_τ, (7C.6) and (7C.4) give

$$\begin{aligned}\langle \mathcal{J}_\tau(\varphi,\pi),\vartheta\rangle &= \int_{\Sigma_\tau} \pi\left(\vartheta_{\mathcal{Y}_\tau}(\varphi)\right)\\ &= \int_{\Sigma_\tau} \pi_A\big(\vartheta^A\circ\varphi - \varphi^A{}_{,i}\vartheta^i\big)\, d^n x_0 \qquad (11\text{A}.1)\end{aligned}$$

for $\vartheta \in \mathfrak{g}_\tau$. But the comments at the end of Section 11C show that this quantity need not vanish whenever $(\varphi,\pi)\in\mathcal{P}_\tau$.

However, let us go back to the infinitesimal equivariance condition (4D.2) which states that $\delta_\xi L = 0$, that is,

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

Suppose $\xi = \xi^\mu\partial_\mu + \xi^A\partial_A$ happens to satisfy

$$\xi^\mu = 0 = \xi^A \quad \text{and} \quad \xi^0{}_{,0} = 0 \qquad (11\text{A}.2)$$

along Y_τ. Then evaluating the infinitesimal equivariance condition on $j^1\phi\circ i_\tau$, where $i_\tau : \Sigma_\tau \to X$ is the inclusion, it collapses to

$$\left[\frac{\partial L}{\partial v^A{}_0}(\xi^A{}_{,0} - v^A{}_i\xi^i{}_{,0})\right]\left(j^1\phi\circ i_\tau\right) = 0;$$

in other words,

$$\pi_A(\xi^A{}_{,0} - \varphi^A{}_{,i}\xi^i{}_{,0}) = 0 \qquad (11\text{A}.3)$$

where $\varphi = \phi\circ i_\tau$.

Now use **A6** to choose an element $\zeta \in \mathfrak{g}$ with $\zeta_X \pitchfork \Sigma_\tau$. We may suppose that bundle coordinates on Y are chosen such that $\zeta_Y \,|\, Y_\tau = \partial/\partial x^0$, so (11A.3) may be rewritten

$$\pi_A \left([\zeta, \xi]^A - \varphi^A{}_{,i} [\zeta, \xi]^i \right) = 0. \tag{11A.4}$$

As $\xi \in \mathfrak{g}$ varies subject to (11A.2), (11A.4)—when integrated over a Cauchy surface—imposes restrictions on the fields $(\varphi, \pi) \in T^*\mathcal{Y}_\tau$. These restrictions are *primary constraints* since (11A.4) holds on the entire image of the Legendre transform. If $(\varphi, \pi) \in \mathcal{P}_\tau$, so that it satisfies (11A.4) , then (11A.1) implies that $\langle \mathcal{J}_\tau(\varphi, \pi), [\zeta, \xi] \rangle = 0$.

We next formalize these observations, with an eye to determining $\mathcal{P}_\tau$ by group-theoretic means.

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

Let $\mathcal{G}$ act on Y by bundle automorphisms. Define

$$\mathfrak{p}_\tau = \{ \xi \in \mathfrak{g}_\tau \mid \xi_{Y_\tau} = 0 \quad \text{and} \quad [\xi, \mathfrak{g}] \subset \mathfrak{g}_\tau \}.$$

In bundle coordinates adapted to Y_τ, elements $\xi \in \mathfrak{p}_\tau$ satisfy (11A.2) along Y_τ (which is the reason we make this particular definition). The subspace $\mathfrak{p}_\tau$ is readily verified to be a Lie subalgebra of $\mathfrak{g}_\tau$; the only nontrivial part to check is closure under bracket. For this, let $\xi, \chi \in \mathfrak{p}_\tau$. Then[29]

$$[\xi, \chi]_{Y_\tau} = -[\xi_{Y_\tau}, \chi_{Y_\tau}] = 0.$$

For $\zeta \in \mathfrak{g}$, the Jacobi identity gives

$$[[\xi, \chi], \zeta] = [\xi, [\chi, \zeta]] - [\chi, [\xi, \zeta]].$$

Then the fact that $\mathfrak{g}_\tau$ is a Lie algebra imply that each term on the right hand side of this expression belongs to $\mathfrak{g}_\tau$.

We assume, as in **A6**, the existence of a $\mathcal{G}$-slicing of Y with generating vector field $\zeta \in \mathfrak{g}$ for which Y_τ is a slice. Throughout this chapter, we will routinely compute in adapted bundle coordinates for which $\zeta_Y = \partial/\partial x^0$ on a neighborhood of Y_τ in Y.

Now define a distribution on Y_τ by

$$\dot{\mathfrak{g}}_\tau = \operatorname{span}_{\mathcal{F}(Y_\tau)} \{ \, [\zeta, \xi]_{Y_\tau} \mid \xi \in \mathfrak{p}_\tau \}. \tag{11B.1}$$

[29] Recall that for left actions, infinitesimal generators satisfy $[\xi, \chi]_M = -[\xi_M, \chi_M]$.

We assume that $\dot{\mathfrak{g}}_\tau$ is regular distribution in the sense that the pointwise evaluation of (11B.1) defines a subbundle of TY_τ. In adapted bundle coordinates, we have

$$[\zeta, \xi]_{Y_\tau} = -\xi^i{}_{,0}\partial_i - \xi^A{}_{,0}\partial_A.$$

Thinking of ζ_Y as a "time direction" then, as both this expression and the notation suggest, $\dot{\mathfrak{g}}_\tau$ *consists of the time derivatives of those infinitesimal generators in* $\mathfrak{g}_\tau$ *which act trivially on* Y_τ.

The following properties of $\dot{\mathfrak{g}}_\tau$ are key.

Proposition 11B.1. *The distribution* $\dot{\mathfrak{g}}_\tau$ *is*

(*i*) *independent of the choice of slicing, and*

(*ii*) *involutive.*

Proof. (*i*) Choose any other $\mathcal{G}$-slicing of Y with generator $\vartheta \in \mathfrak{g}$, and for which Y_τ is a slice. A straightforward calculation using (11A.2) gives

$$[\vartheta_Y, \xi_Y] \,|\, Y_\tau = (\vartheta^0[\zeta_Y, \xi_Y]) \,|\, Y_\tau.$$

The result now follows from the fact that $\vartheta^0 \neq 0$ along Y_τ.

(*ii*) Let $\xi, \chi \in \mathfrak{p}_\tau$. Repeatedly using the Jacobi identity gives

$$2[[\zeta, \xi], [\zeta, \chi]] = [\zeta, [\zeta, [\xi, \chi]]] + [\chi, [\zeta, [\zeta, \xi]]] - [\xi, [\zeta, [\zeta, \chi]]]. \tag{11B.2}$$

We first compute the last term on the right hand side of (11B.2) along Y_τ. We have

$$[\zeta, \chi]_Y = -\chi^\mu{}_{,0}\partial_\mu - \chi^A{}_{,0}\partial_A \quad \text{and} \quad [\zeta, [\zeta, \chi]]_Y = \chi^\mu{}_{,00}\partial_\mu + \chi^A{}_{,00}\partial_A.$$

Since $\xi \in \mathfrak{p}_\tau$, it satisfies $\xi_{Y_\tau} = 0$ and $\xi^0{}_{,0} \,|\, Y_\tau = 0$. Thus, we calculate the last term in (11B.2) to be

$$[\xi, [\zeta, [\zeta, \chi]]]_Y \,|\, Y_\tau = \chi^0{}_{,00}(\xi^i{}_{,0}\partial_i + \xi^A{}_{,0}\partial_A) = -\chi^0{}_{,00}[\zeta, \xi]_Y \,|\, Y_\tau.$$

But by the definition of $\mathfrak{p}_\tau$, $[\zeta, \xi]_Y \,|\, Y_\tau = [\zeta, \xi]_{Y_\tau}$. The middle term of (11B.2) is similar. Hence along Y_τ, (11B.2) becomes

$$2[[\zeta, \xi]_{Y_\tau}, [\zeta, \chi]_{Y_\tau}] = [\zeta, [\zeta, [\xi, \chi]]]_Y \,|\, Y_\tau - \xi^0{}_{,00}[\zeta, \chi]_{Y_\tau} + \chi^0{}_{,00}[\zeta, \xi]_{Y_\tau}. \tag{11B.3}$$

Consider the first term on the right hand side of (11B.3). Since the other terms in (11B.3) are tangent to Y_τ, so is this one; thus,

$$[\zeta, [\zeta, [\xi, \chi]]]_Y \mid Y_\tau = [\zeta, [\zeta, [\xi, \chi]]]_{Y_\tau}. \tag{11B.4}$$

Now $[\xi, \chi] \in \mathfrak{p}_\tau$ and as $[\mathfrak{p}_\tau, \mathfrak{g}] \subset \mathfrak{g}_\tau$, $[\zeta, [\xi, \chi]] \in \mathfrak{g}_\tau$. Moreover, by expanding $[\xi, \chi]_{Y_\tau}$ in coordinates, we conclude from (11A.2) and the Leibniz rule that $[\zeta, [\xi, \chi]]_{Y_\tau} = 0$, whence $[\zeta, [\xi, \chi]] \in \mathfrak{p}_\tau$. Combining this observation with (11B.4), (11B.3) shows that the bracket of two elements of $\dot{\mathfrak{g}}_\tau$ is a linear combination over $\mathcal{F}(Y_\tau)$ of elements of $\dot{\mathfrak{g}}_\tau$. ■

Remark 11B.2. Later in §12B we will give a principal bundle-theoretic construction of $\dot{\mathfrak{g}}_\tau$. ♦

Remark 11B.3. We point out that the construction of $\dot{\mathfrak{g}}_\tau$ utilized a $\mathcal{G}$-slicing, even though this distribution does not depend upon the choice thereof. Consequently we effectively restrict attention to parametrized theories (cf. Remark 6A.1). However, the results of this section will carry over in their entirety to background theories, provided $\mathcal{G}$ is *normal* in $\mathrm{Aut}(Y)$. (By the constructions in §6A the slicing generator $\zeta_Y \in \mathrm{aut}(Y)$; normality is needed to guarantee that $[\zeta, \mathfrak{g}] \subset \mathfrak{g}$ so that (11B.1), etc., make sense.) Typically, in background theories $\mathcal{G} \subset \mathrm{Aut}_{\mathrm{Id}}(Y)$, so this is not a severe restriction. ♦

Remark 11B.4. The fact that $\dot{\mathfrak{g}}_\tau$ forms an involutive distribution will play a crucial role in Chapter 12 when we construct the dynamic and atlas fields. ♦

Remark 11B.5. The construction of $\dot{\mathfrak{g}}_\tau$ is reminiscent of that of the distribution $\mathfrak{g}_{\mathcal{C}_\tau}$ in §8C. In both cases we were led to generalize to the notion of distributions to obtain Lie algebra "actions" in the instantaneous formalism. But the analogy is deeper than this; in fact, as we shall see in the remainder of Chapter 11, $\dot{\mathfrak{g}}_\tau$ plays the same role in regard to the first class primary constraints as $\mathfrak{g}_{\mathcal{C}_\tau}$ does in regard to the first class secondaries. ♦

Lie differentiation of the distribution $\dot{\mathfrak{g}}_\tau$ on Y_τ via (7C.4) gives rise to the distribution

$$\mathrm{span}_{\mathcal{F}(\mathcal{Y}_\tau)}\{[\zeta, \xi]_{\mathcal{Y}_\tau} \mid \xi \in \mathfrak{p}_\tau\}$$

on $\mathcal{Y}_\tau$ and then, by cotangent lift, to the distribution

$$\mathrm{span}_{\mathcal{F}(T^*\mathcal{Y}_\tau)}\{[\zeta, \xi]_{T^*\mathcal{Y}_\tau} \mid \xi \in \mathfrak{p}_\tau\}$$

on $T^*\mathcal{Y}_\tau$. When there is no chance of confusion, we will denote these all by the same symbol $\dot{\mathfrak{g}}_\tau$. Each $\dot{\mathfrak{g}}_\tau$ is involutive and so they define foliations on Y_τ, $\mathcal{Y}_\tau$, and $T^*\mathcal{Y}_\tau$, respectively, which we will collectively denote by $\dot{\mathcal{G}}_\tau$. Our final goal in this subsection is to construct a "momentum map" for $\dot{\mathcal{G}}_\tau$ on $T^*\mathcal{Y}_\tau$. Of course, $\dot{\mathcal{G}}_\tau$ is not a group and does not act on $T^*\mathcal{Y}_\tau$, but it does have "orbits," viz. its leaves. It turns out that this is sufficient.

Recall from §7C that the cotangent action of $\mathcal{G}_\tau$ on $T^*\mathcal{Y}_\tau$ has a momentum map $\mathcal{J}_\tau : T^*\mathcal{Y}_\tau \to \mathfrak{g}_\tau^*$ given by

$$\langle \mathcal{J}_\tau(\varphi, \pi), \xi\rangle = \int_{\Sigma_\tau} \pi \cdot \xi_{\mathcal{Y}_\tau}(\varphi).$$

Writing $\dot{\xi} := [\zeta, \xi]$, we analogously define a map $\dot{\mathcal{J}}_\tau : T^*\mathcal{Y}_\tau \to \dot{\mathfrak{g}}_\tau^*$ by

$$\left\langle \dot{\mathcal{J}}_\tau(\varphi,\pi), \sum_\alpha f_\alpha(\dot{\xi}_\alpha)_{T^*\mathcal{Y}_\tau}\right\rangle = \sum_\alpha \int_{\Sigma_\tau} (f_\alpha \circ \varphi)\, \pi \cdot \big((\dot{\xi}_\alpha)_{\mathcal{Y}_\tau}(\varphi)\big). \tag{11B.5}$$

for $\xi_\alpha \in \mathfrak{p}_\tau$ and functions $f_\alpha \in \mathcal{F}(T^*\mathcal{Y}_\tau)$. In adapted bundle charts, we have

$$\begin{aligned}\big\langle \dot{\mathcal{J}}_\tau(\varphi,\pi), f\,\dot{\xi}_{T^*\mathcal{Y}_\tau}\big\rangle &= \int_{\Sigma_\tau} (f\circ\varphi)\,\pi \cdot \big([\zeta,\xi]_Y \circ \varphi - T\varphi \circ [\zeta,\xi]_X\big) \\ &= -\int_{\Sigma_\tau} (f\circ\varphi)\,\pi_A(\xi^A{}_{,0} - \varphi^A{}_{,i}\xi^i{}_{,0})\, d^n x_0 \end{aligned} \tag{11B.6}$$

by virtue of (7C.4) and (11A.2). Equation (11B.6) and the proof of Proposition 11B.1(i) shows that $\dot{\mathcal{J}}_\tau$ so defined is independent of the choice of slicing.

For future reference, we collect here some useful properties of $\dot{\mathcal{J}}_\tau$.

Proposition 11B.6. (i) *For each $\xi \in \mathfrak{p}_\tau$, we have*

$$\dot{\xi}_{T^*\mathcal{Y}_\tau} \mathbin{\lrcorner} \omega_{T^*\mathcal{Y}_\tau} = \mathbf{d}\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau}\rangle. \tag{11B.7}$$

(ii) *Let $^\vdash$ denote the polar with respect to the canonical symplectic structure on $T^*\mathcal{Y}_\tau$. Then*

$$(\ker T\dot{\mathcal{J}}_\tau)^\vdash = \dot{\mathfrak{g}}_\tau.$$

(iii) *$\dot{\mathcal{J}}_\tau^{-1}(0)$ is coisotropic in $T^*\mathcal{Y}_\tau$.*

Proof. (i) This follows from the observation that $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau}\rangle = \langle \mathcal{J}_\tau, [\zeta,\xi]\rangle$ upon recalling that $\mathcal{J}_\tau$ is a genuine momentum map.

(ii) For each $\xi \in \mathfrak{p}_\tau$ and $\mathcal{V} \in T_{(\varphi,\pi)}(T^*\mathcal{Y}_\tau)$, (11B.7) yields

$$\omega_{T^*\mathcal{Y}_\tau}(\dot{\xi}_{T^*\mathcal{Y}_\tau}(\varphi,\pi), \mathcal{V}) = \mathbf{d}\langle \dot{\mathcal{J}}_\tau(\varphi,\pi), \dot{\xi}_{T^*\mathcal{Y}_\tau}(\varphi,\pi)\rangle \cdot \mathcal{V} \tag{11B.8}$$

$$= \langle T_{(\varphi,\pi)}\dot{\mathcal{J}}_\tau \cdot \mathcal{V}, \dot{\xi}_{T^*\mathcal{Y}_\tau}(\varphi,\pi)\rangle. \tag{11B.9}$$

Thus, $\mathcal{V} \in \ker T_{(\varphi,\pi)}\dot{\mathcal{J}}_\tau$ iff $\omega_{T^*\mathcal{Y}_\tau}(\dot{\xi}_{T^*\mathcal{Y}_\tau}(\varphi,\pi), \mathcal{V}) = 0$ for all $\xi \in \mathfrak{p}_\tau$.

(iii) From (11B.5) we see that $\dot{\mathcal{J}}_\tau^{-1}(0)$ is the annihilator, in $T^*\mathcal{Y}_\tau$, of the distribution $\dot{\mathfrak{g}}_\tau$ on $\mathcal{Y}_\tau$. But it is straightforward to check that the annihilator of a distribution is always coisotropic. ■

In view of these properties it is natural to regard $\dot{\mathcal{J}}_\tau : T^*\mathcal{Y}_\tau \to \dot{\mathfrak{g}}_\tau^*$ as a *momentum map for the foliation* $\dot{\mathcal{G}}_\tau$. It is important to realize, however, that $\dot{\mathcal{J}}_\tau$ is not the momentum map $\mathcal{J}_\tau$ restricted to $\mathfrak{p}_\tau \subset \mathfrak{g}_\tau$, but rather is the "time derivative" of this restriction.

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

Our first main result relates the τ-primary constraint set $\mathcal{P}_\tau \subset T^*\mathcal{Y}_\tau$ with the zero level of the momentum map $\dot{\mathcal{J}}_\tau$ for the foliation $\dot{\mathcal{G}}_\tau$ on $T^*\mathcal{Y}_\tau$. Roughly speaking, the covariance assumption **A1** implies that $\mathcal{L}$ is independent of (certain combinations of) the time derivatives of various field components φ^A. Therefore (certain combinations (11A.4) of) the corresponding conjugate momenta π_A must vanish; these are exactly the momenta $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau}\rangle$ for $\xi \in \mathfrak{p}_\tau$. The conditions $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau}\rangle = 0$ are thus primary constraints. More precisely, we have:

Theorem 11C.1. *Suppose the Lagrangian density is* $\mathcal{G}$*-equivariant. Then*

$$\mathcal{P}_\tau \subset \dot{\mathcal{J}}_\tau^{-1}(0).$$

The detailed proof is contained in the discussion in §11A.

Our next goal is to prove that the constraints $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau}\rangle = 0$ are first class.

Proposition 11C.2. *Assume* **A1**–**A4**. *Then the components* $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau}\rangle$ *of the momentum map for the foliation* $\dot{\mathcal{G}}_\tau$ *of* $T^*\mathcal{Y}_\tau$ *are first class functions.*

We require some preliminaries. Define $\dot{\mathfrak{g}}_{\mathcal{P}_\tau} = \dot{\mathfrak{g}}_\tau \,|\, \mathcal{P}_\tau$ and $\dot{\mathfrak{g}}_{\mathcal{C}_\tau} = \dot{\mathfrak{g}}_\tau \,|\, \mathcal{C}_\tau$. From the construction of $\dot{\mathfrak{g}}_\tau$ and Proposition 8C.6, we obtain:

Lemma 11C.3. *If* **A1**–**A4** *hold, the distributions* $\dot{\mathfrak{g}}_{\mathcal{P}_\tau}$ *and* $\dot{\mathfrak{g}}_{\mathcal{C}_\tau}$ *are tangent to* $\mathcal{P}_\tau$ *and* $\mathcal{C}_\tau$*, respectively.*

Proof of Proposition 11C.2 Let $\xi \in \mathfrak{p}_\tau$ and evaluate (11B.7) along the final constraint set $\mathcal{C}_\tau$. By Lemma 11C.3, $\dot{\xi}_{T^*\mathcal{Y}_\tau} | \, \mathcal{C}_\tau$ is tangent to $\mathcal{C}_\tau$, so the constraint functions $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau} \rangle$ satisfy the first class condition (6E.9). ■

In summary: *The components* $\langle \dot{\mathcal{J}}_\tau, \dot{\xi}_{T^*\mathcal{Y}_\tau} \rangle$ *of the momentum map for the foliation* $\dot{\mathcal{G}}_\tau$ *are first class primary initial value constraints.*

We can say more. As $\dot{\mathfrak{g}}_{\mathcal{P}_\tau} \subset \mathfrak{X}(\mathcal{P}_\tau)$, we can pull (11B.7) back to $\mathcal{P}_\tau$; Theorem 11C.1 gives $\dot{\mathfrak{g}}_{\mathcal{P}_\tau} \subset \ker \omega_{\mathcal{P}_\tau}$ and then Lemma 11C.3 gives

Corollary 11C.4. $\dot{\mathfrak{g}}_{\mathcal{C}_\tau} \subset \ker \omega_{\mathcal{P}_\tau} \cap \mathfrak{X}(\mathcal{C}_\tau)$.

Corollary 11C.4 is the analogue, for primary constraints, of Theorem 8C.2. It proves that the covariant notion of "gauge transformation" is consistent with the corresponding instantaneous notion on the primary level.

Since directions in $\ker \omega_{\mathcal{P}_\tau} \cap \mathfrak{X}(\mathcal{C}_\tau)$ are by definition kinematic (see the discussion surrounding (6E.16)and also at the end of §6E), Corollary 11C.4 proves that *the directions in* $\mathcal{C}_\tau$ *corresponding to* $\dot{\mathfrak{g}}_{\mathcal{C}_\tau}$ *are kinematic.* Thus the evolution of the corresponding (combinations of) fields φ^A is completely arbitrary (or "gauge"). We will see in Chapter 12 that these kinematic fields are closely related to the atlas fields mentioned in the Introduction.

Remark 11C.5. One may wonder why *primary* constraints are correlated with the momentum map for $\dot{\mathcal{G}}_\tau$, whereas only *secondary* constraints are associated with the energy-momentum map for all of $\mathcal{G}$. The reason is that the Legendre transform appears implicitly in $\mathcal{J}^{\mathcal{L}}(\xi) = \mathbb{F}\mathcal{L}^*\mathcal{J}(\xi)$ when working in the Lagrangian representation. (Equivalently, when working in the instantaneous formalism, the energy-momentum map is defined only on $\mathcal{P}_\tau$.) Thus the primary constraints are not detected by the Vanishing Theorem 9A.2. Even so, the primary as well as secondary first class constraints are encoded in the gauge group, but different tools are required to extract them. ♦

Remark 11C.6. While Theorem 11C.1 is the direct analogue of Theorem 10B.1 for primary constraints, the hypotheses in these results are quite different. In particular, the former requires only $\mathcal{G}$-covariance, unlike the latter which also requires localizability, well-posedness, and all fields to be variational. ♦

Examples

In what follows we work in adapted bundle coordinates on Y for which $\zeta_X \,|\, \Sigma_\tau = \partial_0$.

a Particle Mechanics. For the relativistic free particle $\mathfrak{J}_t \equiv 0$ so $\dot{\mathfrak{J}}_t \equiv 0$. Theorem 11C.1 is vacuously true in this case.

b Electromagnetism. We treat the parametrized case first. From (4C.14)–(4C.15), we compute

$$\mathfrak{p}_\tau = \left\{(\xi, \chi) \in \mathfrak{X}(X) \ltimes \mathcal{F}(X) \;\middle|\; j^1\xi \,|\, \Sigma_\tau = 0 \quad \text{and} \quad \mathbf{d}\chi \,|\, \Sigma_\tau = 0\right\}. \tag{11C.1}$$

Then (11B.6) gives

$$\left\langle \dot{\tilde{\mathfrak{J}}}_\tau(A, \mathfrak{E}; g, \pi), (\xi, \chi)^{\cdot}_{T^*\tilde{\mathcal{Y}}_\tau} \right\rangle = -\int_{\Sigma_\tau} \left(\mathfrak{E}^\nu(\chi_{,\nu 0} - A_\mu \xi^\mu{}_{,\nu 0}) - 2\pi^{\sigma\rho} g_{\sigma\mu} \xi^\mu{}_{,\rho 0}\right)\right) d^3x_0$$

where $(\xi, \chi) \in \mathfrak{p}_\tau$. Using (11C.1) to set $\chi_{,i0} = \chi_{,0i} = 0$, etc., this reduces to

$$-\int_{\Sigma_\tau} \left(\mathfrak{E}^0(\chi_{,00} - A_\mu \xi^\mu{}_{,00}) - 2\pi^{\sigma 0} g_{\sigma\mu} \xi^\mu{}_{,00})\right) d^3x_0.$$

Since both $\chi_{,00}$ and $\xi^\mu{}_{,00}$ are arbitrarily specifiable along Σ_τ, we get

$$\dot{\tilde{\mathfrak{J}}}_\tau^{-1}(0) = \{(A, \mathfrak{E}; g, \pi) \mid \mathfrak{E}^0 = 0 \quad \text{and} \quad \pi^{\sigma 0} = 0\}$$

which is a proper subset of $\tilde{\mathcal{P}}_\tau$ which was computed in Example **b** of §6C. Recall also that *ab initio*, $\mathfrak{F}^{\mu\nu}$ is not antisymmetric; it only becomes so when restricted to the covariant primary constraint set, cf. (3B.14).

The computations in the background case are simplified by the absence of the metric momenta $\pi^{\sigma\rho}$. Now we find that $\dot{\mathfrak{J}}_\tau^{-1}(0) = \mathcal{P}_\tau$ exactly, as was computed in Example **b** of §6C. The reason we get equality in the background case but not in the parametrized case will be explained in the next section.

c A Topological Field Theory. Equation (4C.18) yields

$$\mathfrak{p}_\tau = \left\{(\xi, \chi) \in \mathfrak{X}(X) \ltimes \mathcal{F}(X) \;\middle|\; j^1\xi \,|\, \Sigma_\tau = 0 \quad \text{and} \quad \mathbf{d}\chi \,|\, \Sigma_\tau = 0\right\}, \tag{11C.2}$$

and then equation (11B.6) becomes

$$\left\langle \dot{\mathfrak{J}}_\tau(A, \pi), (\xi, \chi)^{\cdot}_{T^*\mathcal{Y}_\tau} \right\rangle = -\int_{\Sigma_\tau} \pi^0(\chi_{,00} - A_\mu \xi^\mu{}_{,00})\, d^2x_0.$$

So

$$\dot{\mathcal{J}}_\tau^{-1}(0) = \{(A, \pi) \mid \pi^0 = 0\}$$

and from Example **c** of §6C we see that Theorem 11C.1 is indeed satisfied.

d Bosonic Strings. Equations (4C.22)–(4C.24) lead to

$$\mathfrak{p}_\tau = \left\{(\xi, \lambda) \in \mathfrak{X}(X) \ltimes \mathcal{F}(X) \;\middle|\; j^1\xi \,|\, \Sigma_\tau = 0 \quad \text{and} \quad \lambda \,|\, \Sigma_\tau = 0\right\}. \tag{11C.3}$$

Now

$$\left\langle \dot{\mathcal{J}}_\tau(\varphi, h, \pi, \varpi), (\xi, \lambda)_{T^*\mathcal{Y}_\tau} \right\rangle = -\int_{\Sigma_\tau} \varpi^{\sigma\rho}\big(2\lambda_{,0} h_{\sigma\rho} - (h_{\sigma\mu}\xi^\mu{}_{,\rho 0} + h_{\rho\mu}\xi^\mu{}_{,\sigma 0})\big)\, d^1x_0$$

$$= -2\int_{\Sigma_\tau} \big(\lambda_{,0}(h_{\sigma\rho}\varpi^{\sigma\rho}) - (h_{\sigma\mu}\varpi^{\sigma 0})\xi^\mu{}_{,00}\big)\, d^1x_0$$

As $\lambda_{,0}$ and $\xi^\mu{}_{,00}$ are arbitrarily specifiable along Σ_τ, the vanishing of $\dot{\mathcal{J}}_\tau$ implies that

$$h_{\sigma\rho}\varpi^{\sigma\rho} = 0 \quad \text{and} \quad h_{\sigma\mu}\varpi^{\sigma 0} = 0. \tag{11C.4}$$

View these as a system of three linear homogeneous equations for the three independent $\varpi^{\sigma\rho}$. Writing these equations out, we compute the determinant of the coefficient matrix to be $h_{11}(\det h_{\sigma\rho})$. Now of course $\det h_{\sigma\rho} \neq 0$ as h is a metric on X. But Σ_τ is assumed to be a spacelike curve, so that $i_\tau^* h = h_{11} dx^1 \otimes dx^1$ must be a positive-definite metric on Σ_τ. (See Remark 6A.4 and Example 6C.**d**.) Thus h_{11} never vanishes on Σ_τ, and it follows that (11C.4) forces $\varpi^{\sigma\rho} = 0$. From Example 6C.**d** we conclude that $\mathcal{P}_\tau = \dot{\mathcal{J}}_\tau^{-1}(0)$.

It is interesting to see what would happen if instead of $\dot{\mathcal{J}}_\tau$ we considered the momentum map $\mathcal{J}_\tau$ for $\mathcal{G}_\tau$. For electromagnetism, (7D.8) gives

$$\langle \mathcal{J}_\tau(A, \mathfrak{E}), \chi \rangle = -\int_{\Sigma_\tau} \mathfrak{E}^\nu \chi_{,\nu}\, d^3x_0$$

$$= \int_{\Sigma_\tau} \mathfrak{E}^0 \chi_{,0}\, d^3x_0 - \int_{\Sigma_\tau} \mathfrak{E}^i{}_{,i}\chi\, d^3x_0.$$

Demanding that this vanish for all $\chi \in \mathcal{F}(X)$ would yield the correct primary constraint $\mathfrak{E}^0 = 0$ but also forces the divergence of the (spatial) electric field to be zero which, however, is a secondary constraint. Analogously, one recovers both the primary constraint $\pi^0 = 0$ and the secondary spatial flatness constraint in this manner in Chern–Simons theory, as well as all constraints—except for the superhamiltonian constraint—for the bosonic string. In particular we see that $\mathcal{P}_\tau$ is not necessarily contained in $\mathcal{J}_\tau^{-1}(0)$.

That the vanishing of $\mathcal{J}_\tau$ yields secondary constraints is not surprising in view of Corollary 7D.2(ii), and that it yields primary constraints in addition is a consequence of the fact that $\dot{\mathcal{J}}_\tau(\dot{\xi}) = \mathcal{J}_\tau([\zeta, \xi])$ for $\xi \in \mathfrak{p}_\tau$. Thus we can assert that

$$\mathcal{E}_\tau^{-1}(0) \subset \mathcal{J}_\tau^{-1}(0) \cap \mathcal{P}_\tau \qquad \text{and} \qquad \mathcal{J}_\tau^{-1}(0) \subset \dot{\mathcal{J}}_\tau^{-1}(0).$$

So might it be possible to short-circuit Chapters 10 and 11 by demanding simply that the momentum map $\mathcal{J}_\tau$ vanish? Unfortunately, no. One obvious reason is that the first inclusion in the above is typically strict, as has been repeatedly emphasized. An equally important reason is that there is no apparent theoretical basis for insisting that $\mathcal{J}_\tau = 0$, such as the Vanishing Theorem for $\mathcal{E}_\tau$ and the covariance assumption **A1** for $\dot{\mathcal{J}}_\tau$. We can conclude *a posteriori* that the vanishing of $\mathcal{J}_\tau$ must yield constraints, but it need not yield all of them.

11D First Class Primary Constraints

In Examples **d** and **e** and the background version of **b** of the previous section, the momentum map $\dot{\mathcal{J}}_\tau$ for the foliation $\dot{\mathcal{G}}_\tau$ on $T^*\mathcal{Y}_\tau$ contains "all" the first class primary constraints present in the theory. We now show that this is true in general provided the gauge group acts "effectively," and prove equality in Theorem 11C.1 under the assumption that all primary constraints are first class.

Corollary 11C.4 shows that $\dot{\mathfrak{g}}_\tau$ consists of gauge generators in the sense of constraint theory. On the other hand, from §6E we know that the distribution $\ker \omega_{\mathcal{P}_\tau} \cap \mathfrak{X}(\mathcal{C}_\tau)$ is locally spanned by the Hamiltonian vector fields of first class primary constraints. Thus, if the components of $\dot{\mathcal{J}}_\tau$ are to comprise *all* such constraints, we must have equality in Corollary 11C.4. We say that the gauge group is ***effective*** if this is the case, and formalize this requirement as an assumption:

A7 Effectiveness. $\dot{\mathfrak{g}}_{\mathcal{C}_\tau} = \ker \omega_{\mathcal{P}_\tau} \cap \mathfrak{X}(\mathcal{C}_\tau)$.

This assumption is to primary constraints as the fullness assumption **A5** is to secondary constraints. It may be viewed as a way of verifying that the gauge group is "large enough," as discussed in §8C. But it may happen that $\mathcal{G}$ does *not* act effectively, even when $\mathcal{G}$ is the correct gauge group, as in the cases of the relativistic free particle or the Nambu string. In such theories the gauge group is not properly synchronized with the geometry, and this suggests that the basic setup of such examples should be modified. We discuss this in some detail in Example **a** following, and at the end of this section.

We now prove a converse to Theorem 11C.1; to do this, we impose certain regularity conditions. The system is required to be almost regular as set forth in **A2** so that, in particular, $\ker \omega_{\mathcal{P}_\tau}$ is a regular distribution on $\mathcal{P}_\tau$. In addition, we shall require that $\dot{\mathfrak{g}}_\tau$ be a regular distribution on $T^*\mathcal{Y}_\tau$. Lemma 11C.3 then implies that $\dot{\mathfrak{g}}_{\mathcal{P}_\tau}$ is regular as well. Finally, we suppose that $\dot{\mathcal{J}}_\tau^{-1}(0)$ is smooth.

Theorem 11D.1. *Suppose that the above assumptions hold. If $\mathcal{G}$ acts effectively, all primary constraints are first class,*[30] *and $\dot{\mathcal{J}}_\tau^{-1}(0)$ is connected, then*

$$\mathcal{P}_\tau = \dot{\mathcal{J}}_\tau^{-1}(0). \tag{11D.1}$$

Proof. Let $^\vdash$ denote the polar with respect to the canonical symplectic structure on $T^*\mathcal{Y}_\tau$. From Proposition 6E.8(*ii*) applied to $\mathcal{P}_\tau \subset T^*\mathcal{Y}_\tau$, we see that $T\mathcal{P}_\tau^\vdash$ is pointwise spanned by the Hamiltonian vector fields of the primary constraints. From §6E, the Hamiltonian vector fields of the first class primary constraints span $\ker \omega_{\mathcal{P}_\tau} \cap T\mathcal{C}_\tau$ pointwise along $\mathcal{C}_\tau$. By assumption all primary constraints are first class, so it follows that

$$T\mathcal{P}_\tau^\vdash \,|\, \mathcal{C}_\tau = \ker \omega_{\mathcal{P}_\tau} \cap T\mathcal{C}_\tau \tag{11D.2}$$

whence

$$T\mathcal{P}_\tau^\vdash \,|\, \mathcal{C}_\tau \subset T\mathcal{C}_\tau \subset T\mathcal{P}_\tau \,|\, \mathcal{C}_\tau.$$

Thus $\mathcal{P}_\tau$ is coisotropic along $\mathcal{C}_\tau$, and the regularity of $\ker \omega_{\mathcal{P}_\tau} = T\mathcal{P}_\tau \cap T\mathcal{P}_\tau^\vdash$ then guarantees that $\mathcal{P}_\tau$ is *globally* coisotropic in $T^*\mathcal{Y}_\tau$.

On the other hand, (11D.2) shows that $\ker \omega_{\mathcal{P}_\tau} \cap T\mathcal{C}_\tau = \ker \omega_{\mathcal{P}_\tau} \,|\, \mathcal{C}_\tau$. By **A7** we conclude that $\dot{\mathfrak{g}}_{\mathcal{C}_\tau} = \ker \omega_{\mathcal{P}_\tau} \,|\, \mathcal{C}_\tau$. This, the fact that $\dot{\mathfrak{g}}_{\mathcal{P}_\tau} \subset \ker \omega_{\mathcal{P}_\tau}$ and the regularity of $\dot{\mathfrak{g}}_{\mathcal{P}_\tau}$ imply that $\dot{\mathfrak{g}}_{\mathcal{P}_\tau} = \ker \omega_{\mathcal{P}_\tau}$. Since $\mathcal{P}_\tau$ is coisotropic, this is equivalent to

$$\dot{\mathfrak{g}}_{\mathcal{P}_\tau} = \mathfrak{X}(\mathcal{P}_\tau)^\vdash. \tag{11D.3}$$

Equation (11B.7) and Theorem 11C.1 imply that

$$\omega_{T^*\mathcal{Y}_\tau}\big(\dot{\mathfrak{g}}_\tau, \mathfrak{X}(\dot{\mathcal{J}}_\tau^{-1}(0))\big) = 0$$

[30] In this context, a primary constraint f is "first class" provided $T\mathcal{C}_\tau^\vdash[f] = 0$.

along $\mathcal{P}_\tau$, whence

$$\mathfrak{X}\big(\dot{\mathfrak{J}}_\tau^{-1}(0)\big)\,|\,\mathcal{P}_\tau \subset \dot{\mathfrak{g}}_\tau^{\vdash}\,|\,\mathcal{P}_\tau = \dot{\mathfrak{g}}_{\mathcal{P}_\tau}^{\vdash}.$$

Taking polars in (11D.3) and substituting in this last result, we get

$$\mathfrak{X}\big(\dot{\mathfrak{J}}_\tau^{-1}(0)\big)\,|\,\mathcal{P}_\tau \subset \mathfrak{X}(\mathcal{P}_\tau).$$

Since $\mathcal{P}_\tau \subset \dot{\mathfrak{J}}_\tau^{-1}(0)$, we obtain

$$T\mathcal{P}_\tau = T\dot{\mathfrak{J}}_\tau^{-1}(0)\,|\,\mathcal{P}_\tau.$$

From the inverse function theorem $\mathcal{P}_\tau$ is an open submanifold of $\dot{\mathfrak{J}}_\tau^{-1}(0)$. Since $\mathcal{P}_\tau$ is closed in $\dot{\mathfrak{J}}_\tau^{-1}(0)$ by **A2**, the connectedness of $\dot{\mathfrak{J}}_\tau^{-1}(0)$ implies that $\mathcal{P}_\tau$ must be all of $\dot{\mathfrak{J}}_\tau^{-1}(0)$. ■

Corollary 11D.2. *With the same assumptions as in Theorem 11D.1, zero is a weakly regular value of* $\dot{\mathfrak{J}}_\tau$.

Proof. Taking polars in (11D.3) results in $\mathfrak{g}_{\mathcal{P}_\tau}^{\vdash} = T\mathcal{P}_\tau$, while Theorem 11D.1 states that $T\mathcal{P}_\tau = T\dot{\mathfrak{J}}_\tau^{-1}(0)$. Combining these results with the polar of Proposition 11B.6(*ii*), we obtain $\ker T\dot{\mathfrak{J}}_\tau = T\dot{\mathfrak{J}}_\tau^{-1}(0)$. ■

We observe that Remark 10B.3 holds in the case of primary constraints as well.

In view of Corollary 11D.2, we have the following partial converse to Theorem 11D.1.

Proposition 11D.3. *If zero is a weakly regular value of* $\dot{\mathfrak{J}}_\tau$ *and* $\mathcal{P}_\tau = \dot{\mathfrak{J}}_\tau^{-1}(0)$, *then* $\mathcal{G}$ *acts effectively.*

Proof. We must show that $\mathfrak{g}_{\mathcal{C}_\tau} = \ker\omega_{\mathcal{P}_\tau} \cap T\mathcal{C}_\tau = T\mathcal{P}_\tau^{\vdash} \cap T\mathcal{C}_\tau$.

Now equation (11B.7) and weak regularity imply that $\dot{\mathfrak{g}}_\tau^{\vdash} = \ker T\dot{\mathfrak{J}}_\tau = T\dot{\mathfrak{J}}_\tau^{-1}(0)$. By assumption we thus have $\dot{\mathfrak{g}}_\tau^{\vdash} = T\mathcal{P}_\tau$ so that $\dot{\mathfrak{g}}_\tau = T\mathcal{P}_\tau^{\vdash}$. But Lemma 11C.3 shows that $\dot{\mathfrak{g}}_{\mathcal{C}_\tau} = \dot{\mathfrak{g}}_\tau\,|\,\mathcal{C}_\tau \subset T\mathcal{C}_\tau$, and the desired result follows. ■

This shows that effectiveness is a necessary condition for the equality of $\mathcal{P}_\tau$ with $\dot{\mathfrak{J}}_\tau^{-1}(0)$. It is not sufficient, however; see Example **c** below.

Remark 11D.4. If $\dot{\mathfrak{J}}_\tau^{-1}(0)$ is not connected, we may still conclude that $\mathcal{P}_\tau$ is a component of $\dot{\mathfrak{J}}_\tau^{-1}(0)$.

Remark 11D.5. Although the regularity assumptions in this Theorem are stringent, they invariably hold in cases of interest (and in particular in Examples **a**–**d**). One finds in practice that both $\mathcal{P}_\tau$ and $\dot{\mathfrak{J}}_\tau^{-1}(0)$ are well-behaved objects; this is a reflection of the fact that primary constraints are usually "simple." (If only the same could be said for secondary constraints!)

Remark 11D.6. Our assumption that all primary constraints are first class masks several subtle points, concerning its implication that $\mathcal{P}_\tau$ is coisotropic in $T^*\mathcal{Y}_\tau$ (cf. the first paragraph in the proof of Theorem 11D.1).

The first point is that $\dot{\mathfrak{J}}_\tau^{-1}(0)$ is *always* coisotropic by Proposition 11B.6(*iii*). Consequently, $\mathcal{P}_\tau$ *must* be coisotropic if (11D.1) is to hold. It would not suffice, however, simply to replace this primary first class assumption by the assumption that $\mathcal{P}_\tau$ is coisotropic in $T^*\mathcal{Y}_\tau$. The reason is that $\mathcal{P}_\tau$ being coisotropic in and of itself does *not* guarantee that all primary constraints are first class. (Whether a (primary) constraint is first class is determined by the geometry of the *final* constraint set $\mathcal{C}_\tau$, not the primary constraint set $\mathcal{P}_\tau$.)

Another point is that the regularity condition on $\ker \omega_{\mathcal{P}_\tau}$ is crucial, since the primary constraints being first class need *not* imply that $\mathcal{P}_\tau$ is coisotropic. The reason is that a constraint being first class restricts its behavior only along the final constraint set $\mathcal{C}_\tau$; in the absence of regularity, the first class condition cannot regulate its behavior elsewhere in $\mathcal{P}_\tau$.

Remark 11D.7. Because $\dot{\mathfrak{J}}_\tau^{-1}(0)$ is always coisotropic in $T^*\mathcal{Y}_\tau$, there is no analogue of Version II of the Energy-Momentum Theorem for primary constraints. ♦

Theorem 11C.1 and Proposition 11C.2 show that the components of $\dot{\mathfrak{J}}_\tau$ are first class primary constraints. Because of Theorem 11D.1, we expect that $\dot{\mathfrak{J}}_\tau$ gives *all* such constraints. We now make this more precise.

Let $I(\mathcal{P}_\tau)$ denote the ideal in $\mathcal{F}(T^*\mathcal{Y}_\tau)$ consisting of primary constraints (i.e., smooth functions on $T^*\mathcal{Y}_\tau$ vanishing on $\mathcal{P}_\tau$). As in §10C, utilizing Theorem 11D.1, Corollary 11D.2, and Remark 10B.3, we may prove:

Proposition 11D.8. *Suppose that* (11D.1) *holds and let zero be a regular value of* $\dot{\mathfrak{J}}_\tau$. *Then*

$$I(\mathcal{P}_\tau) = (\dot{\mathfrak{J}}_\tau).$$

Thus under the stated conditions all primary constraints are first class, and are generated by the components of the momentum map $\dot{\mathfrak{J}}_\tau$. Of course Theorem 11D.1 and Proposition 11D.8 will not hold in the presence of second class primary constraints (in particular, in Example **c**). As in §10C we are able to circumvent this problem using a slight variant of the Dirac bracket construction.

Let $\dot{\mathcal{W}}_\tau \subset T^*Y_\tau$ be defined by the vanishing of the second class (with respect to $\mathcal{C}_\tau$!) primary constraints and $\mathcal{P}_\tau \subset \dot{\mathcal{W}}_\tau$ be defined by the vanishing of the first class primary constraints. Setting $\dot{\mathcal{K}}_\tau = \dot{\mathcal{J}}_\tau \,|\, \dot{\mathcal{W}}_\tau$, we may therefore apply Theorem 11D.1 and Proposition 11D.8 verbatim to $\mathcal{P}_\tau \subset \dot{\mathcal{W}}_\tau$ and $\dot{\mathcal{K}}_\tau$, with the result that

$$I(\mathcal{P}_\tau) = (\dot{\mathcal{K}}_\tau), \tag{11D.4}$$

the ideals now being taken in $\mathcal{F}(\dot{\mathcal{W}}_\tau)$.

Let $I^F(\mathcal{P}_\tau)$ denote the ideal in $\mathcal{F}(T^*\mathcal{Y}_\tau)$ consisting of first class primary constraints, and let $I(\dot{\mathcal{W}}_\tau)$ be the ideal of all smooth functions on $T^*\mathcal{Y}_\tau$ vanishing on $\dot{\mathcal{W}}_\tau$ (so that $I(\dot{\mathcal{W}}_\tau)$ is "generated by the second class primary constraints"). Set $I^F(\dot{\mathcal{W}}_\tau) = I^F(\mathcal{P}_\tau) \cap I(\dot{\mathcal{W}}_\tau)$. Pulling (11D.4) back to $T^*\mathcal{Y}_\tau$, we finally have:

Theorem 11D.9. *Let zero be a regular value of $\dot{\mathcal{J}}_\tau$, and assume regularity as in Theorem 11D.1. If $\mathcal{G}$ acts effectively and $\dot{\mathcal{J}}_\tau^{-1}(0) \cap \dot{\mathcal{W}}_\tau$ is connected, then*

$$I^F(\mathcal{P}_\tau) \equiv (\dot{\mathcal{J}}_\tau) \mod I^F(\dot{\mathcal{W}}_\tau).$$

This result shows that we can recover essentially *all* the first class primary constraints from the momentum map $\dot{\mathcal{J}}_\tau$. The only ones that cannot be obtained in this fashion are the ineffective constraints belonging to $I^F(\mathcal{P}_\tau) \cap I(\dot{\mathcal{W}}_\tau)$; since these first class primaries are "generated by the second class primaries," they are clearly irrelevant insofar as constraint theory is concerned. (This works just as in §10C.)

Examples

In every instance one may verify that both the regularity and connectivity assumptions in Theorem 11D.1 are satisfied. As always, we work in appropriately adapted coordinates.

a Particle Mechanics. Referring to Example **a** of §§6C and 6E, we find that for the relativistic free particle $\dot{\mathfrak{g}}_t = \{0\}$ whereas $\ker \omega_t \cap \mathfrak{X}(\mathcal{C}_t) = \ker \omega_t$ is nontrivial. Thus the time reparametrization group Diff($\mathbb{R}$) does not act effectively and Theorem 11D.1 fails. (This is to be expected from other considerations:

The mass constraint (6C.14) is quadratic in the momenta π_A, and so could not in any case be recovered from $\dot{\mathcal{J}}_t$ which, according to (11B.5), is linear in the π_A.)

To repair this example, we introduce a (positive-definite) metric $h = h_{tt}$ on $\mathbb{R}$, and consider the Lagrangian density

$$\tilde{\mathcal{L}} = -\frac{m}{2}\left(\frac{1}{\sqrt{h}}\|v\|^2 + \sqrt{h}\right) dt \tag{11D.5}$$

viewed as a function also of h and its first jet. Thus we replace Y by $\tilde{Y} = Y \times_X S_2^{0,1}(X)$.[31]

Now the group $\mathcal{G} = \mathrm{Diff}(\mathbb{R})$ of time reparametrizations acts on $\tilde{Y}$ in the obvious way. For $\xi \in \mathcal{F}(\mathbb{R})$, the Lie algebra of $\mathrm{Diff}(\mathbb{R})$, we compute

$$\xi_{\tilde{Y}} = \xi\frac{\partial}{\partial t} - 2h\xi_{,t}\frac{\partial}{\partial h}. \tag{11D.6}$$

Clearly the Lagrangian density (11D.5) is $\mathrm{Diff}(\mathbb{R})$-covariant.

Fix a $\mathcal{G}$-slicing of $\tilde{Y}$ with generating vector field $\zeta_{\tilde{Y}}$ of the form (11D.6) with $\zeta > 0$. The Legendre transform yields

$$\varpi = 0 \tag{11D.7}$$

$$\pi_A = m\zeta^{-1} g_{AB}\dot{q}^B h^{-1/2}, \tag{11D.8}$$

where ϖ is the momentum conjugate to h, along with the Hamiltonian

$$\tilde{H}_{t,\zeta}(q,\pi,h,\varpi) = \frac{1}{2m}\left(g^{AB}\pi_A\pi_B + m^2\right)\zeta\,\sqrt{h} \tag{11D.9}$$

on $\tilde{\mathcal{P}}_t$. There is a single secondary constraint:

$$g^{AB}\pi_A\pi_B + m^2 = 0. \tag{11D.10}$$

Both constraints (11D.7) and (11D.10) are first class.

Notice how this treatment of the relativistic free particle differs thus far from the original one. The Lagrangian density no longer has the square root of (3B.8), but the factor $h^{-1/2}$ compensates for this so that $\tilde{\mathcal{L}}$ is still time reparametrization-equivariant. However, this apparently innocuous modification of the Lagrangian density has dramatic effects insofar as the initial value constraints are concerned. There is a "new" primary constraint $\varpi = 0$, and the

[31] This reformulation of the relativistic free particle has also been employed in Green et al. [1987], §2.1.

"old" primary (6C.14) now turns up as the *secondary* (11D.10). This is possible since the new Hamiltonian (11D.9) does not vanish identically on $\tilde{\mathcal{P}}_t$ (compare (6C.15) with $\zeta^A = 0$). Observe also that the infinitesimal generator (11D.6) has picked up a vertical component.

Returning to our analysis of the primary constraints, (11D.6) gives

$$\tilde{\mathfrak{p}}_t = \{\xi \in \mathcal{F}(X) \mid j^1\xi \mid \Sigma_t = 0\}.$$

But, precisely because of the vertical term in (11D.6), $\dot{\tilde{\mathfrak{g}}}_t$ is no longer trivial; in fact, since

$$[\zeta, \xi]_{\tilde{Y}_t} = 2h\zeta\xi_{,tt}\frac{\partial}{\partial h} \tag{11D.11}$$

we have

$$\dot{\tilde{\mathfrak{g}}}_t = \ \mathrm{span}_{\mathcal{F}(\tilde{Y}_t)}\left\{\frac{\partial}{\partial h}\right\}. \tag{11D.12}$$

On the other hand, from the above,

$$\ker \omega_{\tilde{\mathcal{P}}_t} \cap \mathfrak{X}(\tilde{\mathcal{C}}_t) = \ker \omega_{\tilde{\mathcal{P}}_t} \,|\, \tilde{\mathcal{C}}_t = \mathrm{span}_{\mathcal{F}(\tilde{\mathcal{C}}_t)}\left\{\frac{\delta}{\delta h}\right\}$$

so Diff$(\mathbb{R})$ now acts effectively. Theorem 11D.1 gives

$$\tilde{\mathcal{P}}_t = \{(q, \pi, h, \varpi) \mid \varpi = 0\} = \tilde{\dot{\mathcal{J}}}_t^{-1}(0)$$

as can be verified directly from (11D.11) and (11B.6).

Thus the "Polyakov particle" displays none of the pathologies of the relativistic free particle in its ordinary formulation. The underlying reason this reformulation works is that the offending primary constraint—the mass constraint (6C.14)—is now reincarnated as the secondary constraint (11D.10), and so the problems discussed in the first paragraph of this example are obviated.

Now we show that that the constraint (11D.10) is captured on the secondary level by the energy-momentum map. We compute the latter to be

$$\langle \tilde{\mathcal{E}}_t, \xi \rangle = -\frac{1}{2m}\left(g^{AB}\pi_A\pi_B + m^2\right)\xi\,\sqrt{h} \tag{11D.13}$$

on $\tilde{\mathcal{P}}_t$, whence we have

$$\tilde{\mathcal{E}}_t^{-1}(0) = \{(q, \pi, h) \mid g^{AB}\pi_A\pi_B + m^2 = 0\}. \tag{11D.14}$$

Obviously $\mathcal{E}_t^{-1}(0)$ is coisotropic, so Corollary 10B.6 implies that Diff$(\mathbb{R})$ is full for the Polyakov particle, as can also be verified directly .

In this example Theorem 10B.1 gives a weaker result than Theorem 10B.4 as $\mathcal{C}_\tau$ has two components. $\mathcal{C}_\tau$ is connected only when $m = 0$ in which case it has a conical singularity at the origin. Nonetheless, Proposition 10C.1 remains valid when $m = 0$; see Gotay [1984].

b Electromagnetism. Using (11C.1) and (4C.11)we compute

$$[\zeta, \chi]_{Y_\tau} = -\chi_{,00} \frac{\partial}{\partial A_0}$$

for $\chi \in \mathfrak{p}_\tau$. Thus

$$\dot{\mathfrak{g}}_{\mathcal{P}_\tau} = \operatorname{span}_{\mathcal{F}(\mathcal{P}_\tau)} \left\{ \frac{\delta}{\delta A_0} \right\}.$$

From Example **b** of §6E we see that $\dot{\mathfrak{g}}_{\mathcal{C}_\tau} = \ker \omega_{\mathcal{P}_\tau} \mid \mathcal{C}_\tau = \ker \omega_{\mathcal{P}_\tau} \cap \mathfrak{X}(\mathcal{C}_\tau)$, so that the gauge group $\mathcal{F}(X)$ acts effectively. We have found in Example 11C.**b** that $\dot{\mathcal{J}}_\tau^{-1}(0) = \mathcal{P}_\tau$; from Proposition 11D.8 we conclude that $I(\mathcal{P}_\tau) = (\mathfrak{E}^0)$.

In the parametrized case we have in addition for $\xi \in \mathfrak{p}_\tau$

$$[\zeta, \xi]_{Y_\tau} = 2g_{\sigma\mu}\xi^\mu{}_{,00} \frac{\partial}{\partial g_{\sigma 0}}$$

from (4C.14). So now

$$\dot{\tilde{\mathfrak{g}}}_{\tilde{\mathcal{P}}_\tau} = \operatorname{span}_{\mathcal{F}(\tilde{\mathcal{P}}_\tau)} \left\{ \frac{\delta}{\delta A_0}, \frac{\delta}{\delta g_{\sigma 0}} \right\}$$

which is strictly contained in

$$\ker \tilde{\omega}_{\tilde{\mathcal{P}}_\tau} = \operatorname{span}_{\mathcal{F}(\tilde{\mathcal{P}}_\tau)} \left\{ \frac{\delta}{\delta A_0}, \frac{\delta}{\delta g_{\sigma\rho}} \right\}.$$

Thus parametrized electromagnetism is not effective, and as a consequence the momentum map $\dot{\tilde{\mathcal{J}}}_\tau$ does not capture all the primary constraints in this system as was noted in Example 11C.**b**.

Interestingly, one finds that electromagnetism coupled to ADM gravity is also effective: the appropriate momentum map produces the 5 first class primaries $\mathfrak{E}^0 = 0$ and $\pi^{\sigma 0} = 0$. (In this example there are only 5 primary constraints, and they are all first class.) So at the two extremes—in the background case and coupled to gravity—electromagnetism is effective. It is only in the somewhat artificial intermediate case, viz. when the system contains a nonvariational metric, that this example behaves badly. (Even though it is still producing the "correct"—that is, the first class—primary constraints of the fully dynamic theory.) We have already seen this anomalous behavior on the level of secondary constraints in §10C. The upshot seems to be that *all* fields should be treated variationally.

c A Topological Field Theory. Together (4C.18) and (11C.2) yield

$$[\zeta, (\xi, \chi)]_{Y_\tau} = (A_\mu \xi^\mu{}_{,00} - \chi_{,00})\,\frac{\partial}{\partial A_0}$$

for $(\xi, \chi) \in \mathfrak{p}_\tau$. Thus just as in background electromagnetism

$$\dot{\mathfrak{g}}_{\mathcal{P}_\tau} = \operatorname{span}_{\mathcal{F}(\mathcal{P}_\tau)} \left\{ \frac{\delta}{\delta A_0} \right\} = \ker \omega_{\mathcal{P}_\tau}.$$

Glancing back at Example 6E.**c**, we verify that $\mathrm{Diff}(X) \ltimes \mathcal{F}(X)$ satisfies **A6**. Regardless, Theorem 11D.1 fails in this example, because of the presence of the second class primaries $\pi^i - \epsilon^{0ij} A_j = 0$. Indeed, from Example **c** in the previous section, $\dot{\mathcal{J}}_\tau$ correctly incorporates the first class primary $\pi^0 = 0$, but of course does not "see" these second class primaries.

We take the corresponding Dirac manifold to be

$$\dot{\mathcal{W}}_\tau = \left\{ (A, \pi) \in T^*\mathcal{Y}_\tau \,|\, \pi^i - \epsilon^{0ij} A_j = 0 \right\}$$

with the symplectic structure

$$\dot{\omega}_\tau = \int_{\Sigma_\tau} \left(dA_0 \wedge d\pi^0 + \epsilon^{0ij} dA_i \wedge dA_j \right) \otimes d^2x_0$$

induced by (5C.9). Then $\mathcal{P}_\tau$ given by $\pi^0 = 0$ is coisotropic in $\dot{\mathcal{W}}_\tau$, as expected, and Theorem 11D.9 finally gives

$$I^F(\mathcal{P}_\tau) \equiv (\pi^0) \mod I^F(\dot{\mathcal{W}}_\tau).$$

d Bosonic Strings. For $(\xi, \lambda) \in \mathfrak{p}_\tau$, we have from (4C.22)–(4C.24) and (11C.3) that

$$[\zeta, (\xi, \lambda)]_{Y_\tau} = -2\lambda_{,0} h_{\sigma\rho} \frac{\partial}{\partial h_{\sigma\rho}} + 2h_{\sigma\mu}\xi^\mu{}_{,00} \frac{\partial}{\partial h_{\sigma 0}}.$$

Employing an argument similar to that surrounding (11C.4), we thus have

$$\begin{aligned}
\dot{\mathfrak{g}}_{\mathcal{P}_\tau} &= \operatorname{span}_{\mathcal{F}(\mathcal{P}_\tau)} \left\{ h_{\sigma\rho} \frac{\delta}{\delta h_{\sigma\rho}}, h_{\sigma 0} \frac{\delta}{\delta h_{\sigma 0}} \right\} \\
&= \operatorname{span}_{\mathcal{F}(\mathcal{P}_\tau)} \left\{ \frac{\delta}{\delta h_{\sigma\rho}} \right\} \\
&= \ker \omega_{\mathcal{P}_\tau}.
\end{aligned}$$

Comparing with Example 6E.**d**, we see that the gauge group acts effectively and consequently that $I(\mathcal{P}_\tau) = (\varpi^{\sigma\rho})$.

Remark 11D.10. Consider, for instance, either the relativistic free particle (in its ordinary formulation) or the Nambu string. In both cases, $\mathcal{G} = \mathrm{Diff}(X)$ acts purely "horizontally" on Y (in the sense that $\xi^A = 0$ for all $\xi \in \mathfrak{g}$), and the momentum map $\dot{\mathcal{J}}_\tau$ fails to capture all the first class primary constraints. In either case the problem can be corrected by introducing a metric on X and treating it variationally. This procedure—the "Polyakov trick"—in essence "extends" the action of $\mathcal{G}$ by allowing it to act "vertically" on the metric components. Either system, reformulated in this manner (and with a suitably modified Lagrangian) now has the property that $\mathcal{G}$ acts effectively. Thus, the relativistic free particle and the Nambu string become the Polyakov particle and the Polyakov string, respectively, which, as shown above, behave perfectly well.

The fundamental lesson to be drawn here is that it is often useful to endow the "spacetime" X with a metric; this contention is also supported by our comments in Interlude I and following the definition of Lagrangian slicing in §6A. Furthermore, Example **b** indicates that the metric should be treated variationally as opposed to merely parametrically.

It may be possible to achieve this in many circumstances from first principles according to the following philosophy. Recall the setup for general relativistic fluids or elasticity as sketched in Interlude I. Consider the special case of a fluid in which there is no internal energy, so that the particles all move independently on geodesics. Rather than viewing them one at at time, imagine viewing them as a swarm of particles, with a given initial density. From this point of view one can reformulate the dynamics of a free relativistic particle (our Example **a**), where the base is not $\mathbb{R}$ but spacetime X and fields are maps $\phi : X \to \Sigma \times \mathbb{R}$ where Σ now labels the particle's initial location (or, equivalently, its world line) and ϕ has the interpretation of a ***particle labeling field***. In either case, this point of view turns our Example **a** upside down: what was the base before, namely $\mathbb{R}$, now is part of the fiber, and the base now is standard spacetime. Thus, *the base now comes naturally equipped with a metric which one then treats variationally*, so one might expect effectiveness to hold, as above, without having to resort to the Polyakov trick. This point of view has the pedagogical disadvantage of being overly sophisticated for a free particle, but has the pedagogical advantage of treating elasticity, fluids, and particles in a consistent manner. ♦

Interlude III—Multisymplectic Integrators

IIIA Introduction

A Vision for Integration Algorithms for Field Theories. Multisymplectic integrators are numerical algorithms for field theories that respect the field-theoretic structure at the discrete level. They have very attractive numerical properties, such as getting the total energy as well as the energy balance between modes in the system correct and, in addition, computing the total momentum as well as the multimomentum fluxes exactly on the discrete level. Considerable numerical evidence shows that these are desirable properties indeed and are especially useful in situations in which one would like to get good mechanical properties without a fully resolved simulation. These sorts of integration algorithms have been implemented for nonlinear wave equations and for nonlinear elasticity, for example, and have proven to be very effective and competitive with existing algorithms. Other systems, such as fluids, are actively being developed; in addition, relativistic theories, including general relativity, comprise an important area for future research.

One of the reasons these integrators are promising for general relativity is that if they are compatible with the gauge symmetry structure of a theory, then they automatically preserve, in the multisymplectic sense, the multimomentum mapping and the constraints of the theory—exactly on the discrete level. In general relativity, these constraints are of course the superhamiltonian and supermomentum constraints and it is known numerically that the preservation of these constraints are one of several crucial roadblocks in the development of integrators sufficiently accurate to be able to compute, for example, the wave forms of gravitational waves that are emitted from inspiraling black holes. This is felt to be a critical ingredient in the detection of gravitational waves. See, for instance, Holst et al. [2004].

Not surprisingly, there remain many difficulties to overcome. One problem is how to discretize the diffeomorphism group, the gauge group of general relativity. However, recently, in connection with fluid mechanics, a candidate has arisen, namely the group of doubly stochastic matrices (also known as the *incompressible Markov group*) associated with a given mesh (see Ebin and Marsden [1970], Desbrun, Kanso, Marsden, Mackenzie, and Tong [2006], and Johnson

[1985]. A second problem is the need for development of and incorporation of discrete differential geometry (perhaps in the form of Regge calculus) into the formalism of multisymplectic integrators. This is currently believed to be important for electromagnetism; see for instance, Bossavit [1998].

Despite these roadblocks, the outlook for the continued development of multisymplectic integrators is bright and they have, in fact, already been extremely successful in various contexts, as has been mentioned.

In this Interlude, we will focus on a special class of these integrators known as ***variational integrators***. The key idea here is to not directly discretize the Euler–Lagrange equations, which would be the normal procedure in numerical analysis, but to *discretize the variational principle on which they are based.* As we shall see variational integrators are automatically multisymplectic. In particular, these algorithms preserve the multisymplectic structure in the sense that a discrete version of the *multisymplectic form formula* holds; this is the multisymplectic analog of the fact that the flows of the Euler–Lagrange equations and Hamilton's equations consist of symplectic transformations (see §IIIC). The variational approach allows one additional flexibility in the design of the algorithms and also helps in understanding how one passes from particle mechanics to field theories.

The basic ideas of multisymplectic integration are already present in simple examples, such as particle mechanics, which we discuss after a brief historical aside.

A Brief History and Abbreviated Literature. The history of variational integrators is somewhat complex, but its roots can be traced to Hamilton-Jacobi theory and to discrete optimal control. We shall not recount this early history, but refer to Marsden and West [2001] for an extensive survey, and pick up the story in more recent times. One line of motivation comes from the work of Suris [1989, 1990] and Moser and Veselov [1991]. Here one finds the origins of many of the ideas of discrete mechanics based on discrete variational principles. These ideas were developed and applied to a number of examples, such as particle mechanics and rigid bodies, in Wendlandt and Marsden [1997]. Still in the standard context of Lagrangian and Hamiltonian mechanics, there were several other key developments. in particluar, the incorporation of external forces (such as dissipative forces) as well as the study of the Newmark scheme (one of the workhorses of computational mechanics) was given in Kane, Marsden, Ortiz, and West [2000]. It was shown there that variational methods respect the energy budget of a system in a way that is far superior to standard algorithms.

It should be also noted that there are deep links between the variational method for discrete mechanics and integrable systems. This area of research was started by Moser and Veselov [1991] and was continued by many others, notably by Bobenko and Suris; we refer the reader to the book Suris [2003] for more information. The main and very interesting example studied by Moser and Veselov was to find an integrable discretization of the n-dimensional rigid body, an integrable system; see also Bloch, Crouch, Marsden, and Ratiu [2002] for further insight into the discretization process in this case.

This line of work on mechanical integrators provided a natural avenue of investigation into the field-theoretical case for the simple reason—as is evident in this book—that it is based on variational methods. Again, the idea is to discretize not the field equations, but the variational principle and to show that because of the resulting discrete variational structure, the multisymplectic properties get inherited on the discrete level. The first major paper to take this viewpoint is Marsden et al. [1998]—hereafter abbreviated as MPS. The main numerical example treated there is the sine-Gordon equation; the algorithm was shown to have excellent numerical properties, such as good energy and conservation law behavior. We present this example in §IIID.

The variational viewpoint was further developed and applied to nonlinear elasticity in a series of papers on asynchronous variational integrators in Lew et al. [2003, 2004]. Their methods have also been applied to collision problems (see Fetecau, Marsden, Ortiz, and West [2003] and references therein), to mesh adaptation (see, for instance, Thoutireddy and Ortiz [2004]), and many other problems. The developments in this area continue apace.

Around the same time, a rather different method, based on a "many-symplectic" viewpoint, appeared in the work of Bridges [1997] and subsequent papers. This was also numerically implemented in Bridges and Reich [2001] amongst other works. The two viewpoints are related; see, for example, Marsden and Shkoller [1999], but the developments have proceeded along their own paths. Other (nonvariational) approaches to multisymplectic integrators have been studied by Moore and Reich [2003] and Islas and Schober [2004].

IIIB Basic Ideas of Variational Integrators

The idea of variational integrators is quite simple: one obtains algorithms by forming discrete versions of variational principles. For conservative systems (including field theories) one typically uses Hamilton's principle, while for dissipative or forced systems one uses the Lagrange–d'Alembert principle. For simplicity we limit ourselves to illustrating the former in the case of finite-dimensional systems.

Recall from basic mechanics (see, for example, Marsden and Ratiu [1999]) the *configuration space form of Hamilton's principle*. Let an autonomous mechanical system have an N-dimensional configuration manifold Q and be described by a Lagrangian $L : TQ \to \mathbb{R}$. Then the principle states that the action integral

$$S = \int_a^b L(q(t), \dot{q}(t))\, dt$$

is stationary for curves $q(t)$ in Q with fixed endpoints as in Figure III.1.

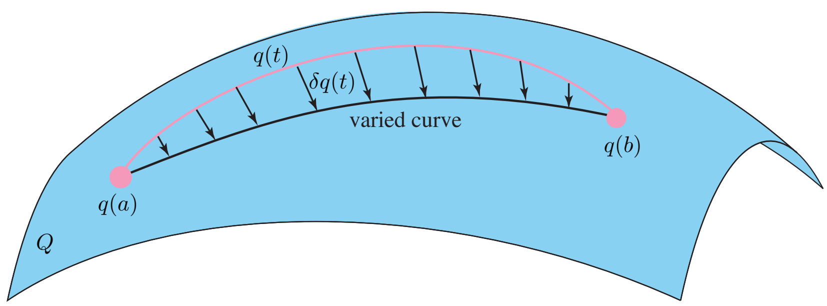


Figure III.1: The configuration space form of Hamilton's principle

Of course, by Theorem 3B.1 Hamilton's principle is equivalent to the Euler–Lagrange equations (3B.7).

In ***discrete mechanics*** from the Lagrangian point of view, one first forms a ***discrete Lagrangian***, a function L_d of two points $q_0, q_1 \in Q$ and a time step h by approximating the the action integral along an exact trajectory with a quadrature rule:

$$L_d(q_0, q_1, h) \approx \int_0^h L\big(q(t), \dot{q}(t)\big)\, dt$$

where $q(t)$ is an *exact* solution of the Euler–Lagrange equations for L joining q_0 to q_1 over the *time step interval* $0 \le t \le h$. Holding h fixed for the moment, we may regard L_d as a mapping $L_d : Q \times Q \to \mathbb{R}$. This way of thinking of the discrete Lagrangian as a function of two nearby points (which take the place of a discrete position and velocity) goes back to the origins of Hamilton–Jacobi theory itself, but appears explicitly in the discrete optimal control literature in the 1960s, and was exploited effectively by, for example, Suris [1990], Moser and Veselov [1991], and Wendlandt and Marsden [1997]. It is a point of view that is crucial for the development of the theory.

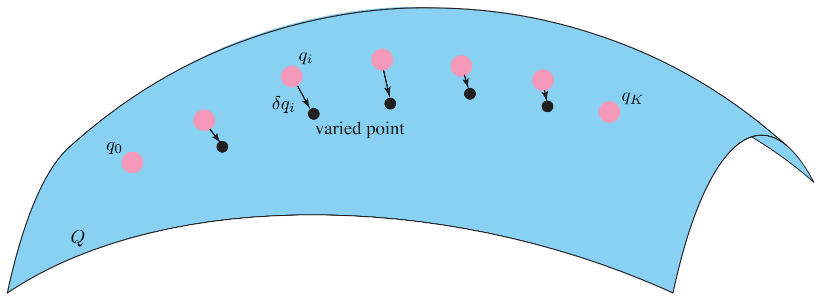


Figure III.2: The discrete form of Hamilton's principle

Given a discrete Lagrangian L_d, the discrete theory proceeds in its own right as follows. Given a sequence $q_0, \ldots, q_K$ of points in Q, form the ***discrete action sum:***

$$S_d = \sum_{k=0}^{K-1} L_d\left(q_k, q_{k+1}, h_k\right).$$

Then the ***discrete Hamilton configuration space principle*** requires us to seek a critical point of S_d with fixed end points, q_0 and q_K. See Figure III.2 Taking the special case of three points q_{i-1}, q_i, q_{i+1}, so the discrete action sum is

$$L_d\left(q_{i-1}, q_i, h_{i-1}\right) + L_d\left(q_i, q_{i+1}, h_i\right),$$

and varying with respect to the middle point q_i gives the ***discrete Euler–Lagrange*** (***DEL***) equations:

$$D_2 L_d\left(q_{i-1}, q_i, h_{i-1}\right) + D_1 L_d\left(q_i, q_{i+1}, h_i\right) = 0. \tag{IIIB.1}$$

One arrives at exactly the same result using the full discrete variational principle. Equation (IIIB.1) defines, perhaps implicitly, the ***DEL algorithm***: $(q_{i-1}, q_i) \mapsto (q_i, q_{i+1})$.

Example: Particle Mechanics. Let $\mathbf{M}$ be a positive definite symmetric 3×3 matrix and $V : \mathbb{R}^3 \to \mathbb{R}$ be a given potential. Choose a discrete Lagrangian on $\mathbb{R}^3 \times \mathbb{R}^3$ of the form

$$L_d(h, \mathbf{q}_0, \mathbf{q}_1) = h\left[\frac{1}{2}\left(\frac{\mathbf{q}_1 - \mathbf{q}_0}{h}\right)^T \mathbf{M}\left(\frac{\mathbf{q}_1 - \mathbf{q}_0}{h}\right) - V(\mathbf{q}_0)\right], \tag{IIIB.2}$$

which arises in an obvious way from its continuous counterpart by using a "rectangle rule" on the action integral. For this discrete Lagrangian, the DEL equations are readily worked out to be

$$\mathbf{M}\left(\frac{\mathbf{q}_{k+1} - 2\mathbf{q}_k + \mathbf{q}_{k-1}}{h^2}\right) = -\nabla V(\mathbf{q}_k),$$

a discretization of Newton's equations, using a simple finite difference rule for the derivative. ♦

In mechanics, the initial conditions are typically specified as a position and a velocity or momentum rather than two positions; therefore it is beneficial to write the DEL equations in a ***position-momentum form***. To this end, define the momentum at time t_k to be:[32]

$$p_k := D_2 L_d(q_{k-1}, q_k) = -D_1 L_d(q_k, q_{k+1})$$

where the second equality holds due to the DEL equations. The position-momentum form of the discrete equations is then given by:

$$p_k = -D_1 L_d(q_k, q_{k+1}) \ , \qquad p_{k+1} = D_2 L_d(q_k, q_{k+1}). \tag{IIIB.3}$$

For (q_k, p_k) known, (IIIB.3)(left) is an (often implicit) equation whose solution gives q_{k+1}. q_{k+1} is then substituted in (IIIB.3)(right) to find p_{k+1}. This provides an update rule in phase space.

One may also approach discrete mechanics from a Hamiltonian point of view; here, *Hamilton's phase space principle* comes into play. Many algorithms, such as the midpoint rule and symplectic Runge–Kutta ("SPARK") schemes, appear more naturally in the Hamiltonian context, as pointed out in Marsden and West [2001]. Additionally, there is a hybrid of the Hamilton configuration and phase space principles known as the *Hamilton-Pontryagin principle* Yoshimura and Marsden [2006], which has counterparts in some field theories and is important in what is called the discontinuous Galerkin method. For instance, in nonlinear elasticity, it is called the "Hu-Washizu principle." It would be of substantial interest to extend this principle to the general context of field theories as given in this book.

It is shown in Kane, Marsden, Ortiz, and West [2000] that the widely used Newmark scheme is variational in our sense. (Structurally, the Newmark scheme is similar to the example presented above. One can argue that the variational

[32] Henceforth we suppress the time step dependence in L_d.

nature of the Newmark scheme is one of the reasons for its excellent performance.) Many other standard integrators are variational as well, including the midpoint rule, SPARK schemes, etc.; we refer to Suris [1990] for details. In fact, *every* symplectic integrator is variational (Marsden and West [2001]).[33]

IIIC Properties of Variational Integrators

We motivate the properties of variational integrators first with a few examples and then we briefly discuss a bit of theory involving momentum maps and symplecticity.

Numerical Motivation. No matter what the choice of the discrete Lagrangian, variational integrators are in the conservative case *momentum conserving* and *symplectic* in a way that will be made precise below. "Momentum conserving" means that when the discrete system has a symmetry, then there is a discrete Noether theorem that gives a quantity that is exactly conserved at the discrete level. Figure III.3(a) illustrates the sort of qualitative difference that momentum conservation gives in solar system dynamics, and in (b) we illustrate in the case of two-dimensional systems that "symplectic" means area preserving, even with large distortions.[34]

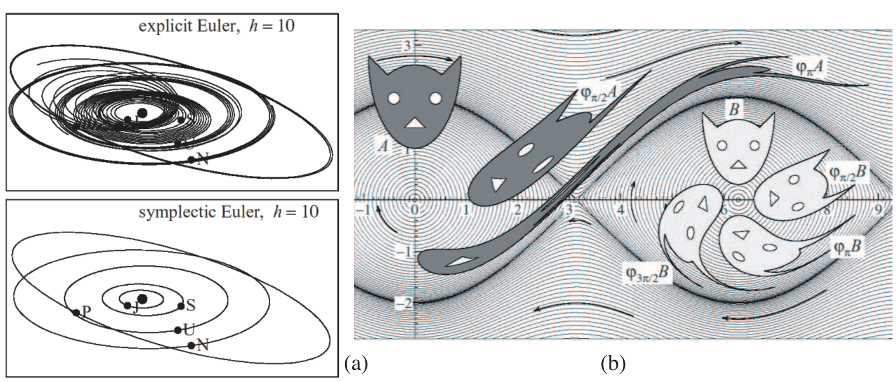
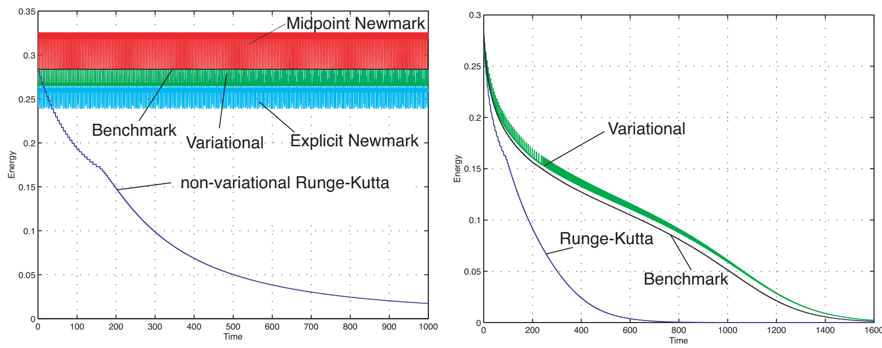


Figure III.3: (a) Variational integrators give good qualitative behavior for comparable computational effort. (b) Variational integrators preserve the symplectic form in phase space (area, in two dimensions).

[33] It is not known if the analogous statement is true for multisymplectic integrators.
[34] These figures are due to Hairer, Lubich, and Wanner [2001].



(a) Conservative mechanics (b) Dissipative mechanics

Figure III.4: Showing the excellent energy behavior for both conservative and dissipative systems: a particle in $\mathbb{R}^2$ with a radially symmetric polynomial potential (left); with small dissipation proportional to the velocity (right).

Variational integrators have remarkably good energy behavior in both the conservative and dissipative cases (for the latter, one discretizes the Lagrange-d'Alembert principle); consider, for example, the system described in Figure III.4, namely a particle moving in the plane. This figure illustrates the fact that variational integrators have long time energy stability (as long as the time step is reasonably small). This is an important property, but it is also a deep one from the theoretical point of view and is observed to hold numerically in many cases when the theory cannot strictly be verified; the key technique is known as *backward error analysis* and it seeks to show that the algorithm is, up to exponentially small errors, the exact time integration of a nearby Hamiltonian system, an idea going back to Neishtadt [1984]. See Hairer and Lubich [2000] for a thorough analysis.

We now turn to a brief outline of some of the theory.

Momentum Conservation. The discrete version of Noether's (first) theorem parallels the continuous case discussed in §4D. Consider a one-parameter family of discrete time curves $\{q_k^\varepsilon\}_{k=0}^K$, with $q_k^0 = q_k$, such that $L_d(q_k^\varepsilon, q_{k+1}^\varepsilon) = L_d(q_k, q_{k+1})$ for all ε and k. The corresponding infinitesimal symmetry is written

$$\xi_k = \left.\frac{\partial q_k^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0}. \tag{IIIC.1}$$

Invariance of the discrete Lagrangian implies invariance of the action sum, and so its ε-derivative will be zero. Assuming that $\{q_k\}$ is a solution trajectory, then variation of the discrete action sum gives

$$0 = \frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \sum_{k=0}^{K-1} L_d(q_k^\varepsilon, q_{k+1}^\varepsilon) = D_1 L_d(q_0, q_1)\cdot\xi_0 + D_2 L_d(q_{K-1}, q_K)\cdot\xi_K. \quad \text{(IIIC.2)}$$

Observing that $0 = D_1 L_d(q_0, q_1) \cdot \xi_0 + D_2 L_d(q_0, q_1) \cdot \xi_1$ as L_d is invariant, we thus have the ***discrete Noether theorem***

$$D_2 L_d(q_{K-1}, q_K) \cdot \xi = D_2 L_d(q_0, q_1) \cdot \xi, \quad \text{(IIIC.3)}$$

where the discrete momentum in the direction ξ is given by $D_2 L_d(q_k, q_{k+1}) \cdot \xi$.

Example: Particle Mechanics. Consider the discrete mechanical Lagrangian (IIIB.2) and assume that V is a function of $\|\mathbf{q}\|$ only. (This is the case of a particle in a radial potential, for example.) Then L_d is invariant under rotations $\mathbf{q}_k^\varepsilon = \exp(\varepsilon\mathbf{\Omega})\mathbf{q}_k$, for any skew-symmetric matrix $\mathbf{\Omega} \in \mathbb{R}^{3\times3}$. Evaluating (IIIC.3) in this case gives

$$\mathbf{q}_K \times \mathbf{M}\left(\frac{\mathbf{q}_K - \mathbf{q}_{K-1}}{t_K - t_{K-1}}\right) = \mathbf{q}_1 \times \mathbf{M}\left(\frac{\mathbf{q}_1 - \mathbf{q}_0}{t_1 - t_0}\right). \quad \text{(IIIC.4)}$$

Thus, we get expressions for the discrete angular momentum, and have shown that it is conserved on the discrete level. Note that while this result might seem obvious, in more complicated examples this will not be the case. ♦

As in the continuous case, we can extend the above derivation to general Lie groups and define a full discrete momentum map $J_{L_d} : Q \times Q \to \mathfrak{g}^*$ by $J_{L_d}(q_0, q_1)\cdot\xi = D_2 L_d(q_0, q_1)\cdot\xi_Q(q_1)$. (In fact there are two discrete momentum maps, corresponding to $D_1 L_d$ and $D_2 L_d$, but they are equal whenever L_d is invariant.) From here one can proceed to develop a theory of *discrete reduction* as in Bobenko and Suris [1999a,b], Marsden, Pekarsky, and Shkoller [1999, 2000], and Jalnapurkar, Leok, Marsden and West [2006].

Symplecticity. In addition to conserving energy and momenta, Lagrangian evolution also preserves the ***Lagrange form*** $\omega_L = -d\theta_L$, where[35]

$$\theta_L = \frac{\partial L}{\partial q^A} dq^A.$$

[35] Intrinsically, θ_L is the pullback to TQ of the mechanical Cartan form (3B.6) on $J^1 Y = \mathbb{R} \times TQ$. The 2-form ω_L, which is the Lagrangian counterpart of the symplectic form, is actually symplectic iff L is regular.

Extended discussions can be found in MPS and Marsden and Ratiu [1999], and the analogous statement for continuous Hamiltonian evolution has already been proved explicitly in Corollary 6D.2. As the symplecticity of continuous time Lagrangian systems is a direct consequence of the variational structure, there is thus an analogous property of discrete Lagrangian systems. We proceed to discover this conservation law "by hand" using the variational structure in the mechanical case.

Consider a two-parameter set of initial conditions $\{(q_0^{\varepsilon,\nu}, q_1^{\varepsilon,\nu})\}$ and let $\{q_k^{\varepsilon,\nu}\}_{k=0}^K$ be the resulting discrete trajectories. We denote the corresponding variations by

$$\delta q_k^\varepsilon = \frac{\partial}{\partial \nu} q_k^{\varepsilon,\nu}\bigg|_{\nu=0} \qquad \delta \bar{q}_k^\nu = \frac{\partial}{\partial \varepsilon} q_k^{\varepsilon,\nu}\bigg|_{\varepsilon=0} \qquad \delta^2 q_k = \frac{\partial^2}{\partial \varepsilon \partial \nu} q_k^{\varepsilon,\nu}\bigg|_{\varepsilon,\nu=0}$$

and we write $\delta q_k = \delta q_k^0$, $\delta \bar{q}_k = \delta \bar{q}_k^0$ and $q_k^\varepsilon = q^{\varepsilon,0}$ for $k = 0, \ldots, K$. The second derivative of the action sum is given by

$$\begin{aligned}
&\frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \frac{\partial}{\partial \nu}\bigg|_{\nu=0} S_d(\{q_k^{\varepsilon,\nu}\}) \\
&\quad= \frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \Big(DS_d(\{q_k^\varepsilon\}) \cdot \delta q^\varepsilon \Big) \\
&\quad= \frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \Big(D_{1i} L_d(q_0^\varepsilon, q_1^\varepsilon)\, (\delta q_0^\varepsilon)^i + D_{2i} L_d(q_{K-1}^\varepsilon, q_K^\varepsilon)\, (\delta q_K^\varepsilon)^i \Big) \\
&\quad= D_{1j} D_{1i} L_d(q_0, q_1) \delta q_0^i \delta \bar{q}_0^j + D_{2j} D_{1i} L_d(q_0, q_1) \delta q_0^i \delta \bar{q}_1^j \\
&\qquad+ D_{1j} D_{2i} L_d(q_{K-1}, q_K) \delta q_K^i \delta \bar{q}_{K-1}^j + D_{2j} D_{2i} L_d(q_{K-1}, q_K) \delta q_K^i \delta \bar{q}_K^j \\
&\qquad+ D_{1i} L_d(q_0, q_1) \delta^2 q_0^i + D_{2i} L_d(q_{K-1}, q_K) \delta^2 q_K^i.
\end{aligned}$$

By symmetry of mixed partial derivatives, reversing the order of differentiation above will give an equivalent expression. Subtracting one from the other will thus give zero, and rearranging the resulting equation we obtain

$$\begin{aligned}
&D_{1j} D_{2i} L_d(q_{K-1}, q_K) \left[\delta q_K^i \delta \bar{q}_{K-1}^j - \delta \bar{q}_K^i \delta q_{K-1}^j \right] \\
&\qquad\qquad = D_{2j} D_{1i} L_d(q_0, q_1) \left[\delta \bar{q}_0^i \delta q_1^j - \delta q_0^i \delta \bar{q}_1^j \right]. \qquad \text{(IIIC.5)}
\end{aligned}$$

Directly from the symmetry of mixed partial derivatives we have

$$D_{2j}D_{1i}L_d(q_0,q_1)\left[\delta\bar{q}_0^i\delta q_1^j - \delta q_0^i\delta\bar{q}_1^j\right] = D_{1j}D_{2i}L_d(q_0,q_1)\left[\delta q_1^i\delta\bar{q}_0^j - \delta\bar{q}_1^i\delta q_0^j\right]. \qquad \text{(IIIC.6)}$$

Substituting this into (IIIC.5) now gives

$$D_{1j}D_{2i}L_d(q_{K-1},q_K)\left[\delta q_K^i\delta\bar{q}_{K-1}^j - \delta\bar{q}_K^i\delta q_{K-1}^j\right] = D_{1j}D_{2i}L_d(q_0,q_1)\left[\delta q_1^i\delta\bar{q}_0^j - \delta\bar{q}_1^i\delta q_0^j\right]. \qquad \text{(IIIC.7)}$$

We can now see that each side of this equation is an antisymmetric bilinear form, which we call the ***discrete Lagrange form***, evaluated on the variations δq_k and $\delta\bar{q}_k$. The two sides give this expression at the first time step and the final time step, so we conclude that the discrete Lagrange form is preserved by the time evolution of the discrete system.

Intrinsically we can identify two one-forms $\theta^+_{L_d} = D_2L_d\,dq_1$ and $\theta^-_{L_d} = D_1L_d\,dq_0$, so that $dS_d = (F^K_{L_d})^*\theta^+_{L_d} + \theta^-_{L_d}$, where $F^K_{L_d}$ is the discrete time flow. Then $0 = d^2S_d = (F^K_{L_d})^*(d\theta^+_{L_d}) + d\theta^-_{L_d}$ and so defining the discrete two-forms $\omega^{\pm}_{L_d} = -d\theta^{\pm}_{L_d}$ gives $(F^K_{L_d})^*\omega^+_{L_d} = -\omega^-_{L_d}$, which is the intrinsic form of (IIIC.5). However, we observe that $0 = d^2L_d = d(\theta^+_{L_d} + \theta^-_{L_d}) = -\omega^+_{L_d} - \omega^-_{L_d}$ and hence $\omega^+_{L_d} = -\omega^-_{L_d}$, which is (IIIC.6). Combining this with our previous expression then gives $(F^K_{L_d})^*\omega^+_{L_d} = \omega^+_{L_d}$ as the intrinsic form of (IIIC.7), which is discrete symplecticity of the evolution.

IIID Multisymplectic and Asynchronous Variational Integrators

One of the beautiful things about the variational approach is that it suggests a natural extension to the PDE case. Namely one should discretize, in spacetime, the variational principle for a given field theory, such as electromagnetism, elasticity, or gravity, etc. To lay the groundwork for such a discretization, one replaces the discrete time points with a mesh in spacetime and replaces points in Q with clusters of points in Y (so that one can represent the needed derivatives of the fields). These clusters of points may be regarded as a discretization of the first jet bundle. (See Figure III.5.) Then the theory proceeds in much the same way as described above.

The basic set-up and feasibility of this idea for infinite-dimensional systems was first demonstrated in MPS. This paper showed that there were discrete field-

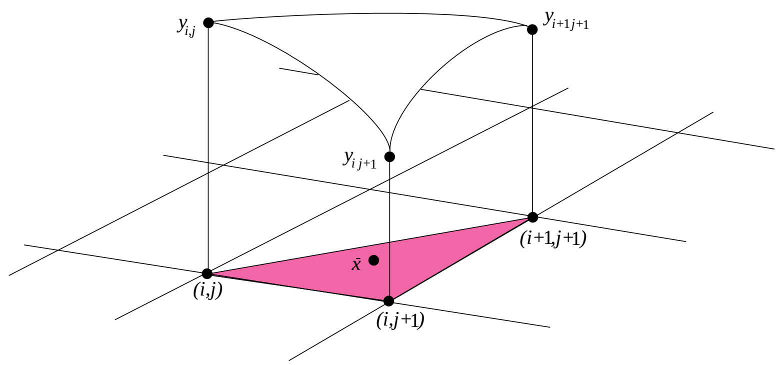


Figure III.5: Depiction of the heuristic interpretation of an element of J^1Y when X is discrete.

theoretic analogs of all the structures one has in mechanics with some obvious modifications; mainly, the symplectic structure gets replaced by the appropriate multisymplectic structure. The key new ingredient is the ***multisymplectic form formula*** which states that the integral of a bilinear expression in the field variations, built out of the multisymplectic form, integrated over the boundary of a region in spacetime and evaluated at a solution, is zero. This generalizes the symplecticity property, in which the difference of the values of the Lagrange form at two ends of a temporal interval is zero. This is proved, both in the continuous and discrete settings in exactly the same way as was done above for mechanics; namely, it follows from the fact that the second differential of the action function restricted to the space of solutions is zero. As in the finite-dimensional case, all of these properties follow from the fact that one has a discrete variational principle.

Example: Nonlinear Wave Equation. Consider the scalar nonlinear wave equation

$$\frac{\partial^2 \phi}{\partial {x^0}^2} - \triangle \phi - N'(\phi) = 0, \tag{IIID.1}$$

where $\triangle$ is the Laplace-Beltrami operator and N is a real-valued C^∞ function of one variable. For concreteness, we fix $n = 1$ and take $X = \mathbb{R}^2$ and $Y = X \times \mathbb{R}$.

Equation (IIID.1) is governed by the Lagrangian density

$$\mathcal{L} = \left\{ \frac{1}{2} \left[\left(\frac{\partial \phi}{\partial x^0} \right)^2 - \left(\frac{\partial \phi}{\partial x^1} \right)^2 \right] + N(\phi) \right\} dx^1 \wedge dx^0.$$

To discretize it, we visualize each triangle $\Delta \subset X$ as having base length h and height k; see Figure III.5. We then think of the continuous jet $j^1\phi$ as corresponding to a discrete jet as follows:

$$\phi(\bar{x}_{ij}) = \frac{y_{ij} + y_{i+1\,j} + y_{i+1\,j+1}}{3},$$

$$\frac{\partial \phi}{\partial x^0}(\bar{x}_{ij}) = \frac{y_{i\,j+1} - y_{ij}}{h}, \quad \frac{\partial \phi}{\partial x^1}(\bar{x}_{ij}) = \frac{y_{i+1\,j+1} - y_{i\,j+1}}{k},$$

where $\bar{x}_{ij}$ is at the center of Δ. This leads to the discrete Lagrangian

$$L = \frac{1}{2}\left(\frac{y_2 - y_1}{h}\right)^2 - \frac{1}{2}\left(\frac{y_3 - y_2}{k}\right)^2 + N\left(\frac{y_1 + y_2 + y_3}{3}\right),$$

where we use y_1, y_2, y_3 as generic labels for the vertices $y_{ij}, y_{i\,j+1}, y_{i+1\,j+1}$ of a triangle in Y covering Δ.

The corresponding DELF equations are

$$\frac{y_{i+1\,j} - 2\,y_{ij} + y_{i-1\,j}}{k^2} - \frac{y_{i\,j+1} - 2\,y_{ij} + y_{i\,j-1}}{h^2} + \frac{1}{3}N'\left(\frac{y_{ij} + y_{i\,j+1} + y_{i+1\,j+1}}{3}\right) + \frac{1}{3}N'\left(\frac{y_{i\,j-1} + y_{ij} + y_{i+1\,j}}{3}\right) + \frac{1}{3}N'\left(\frac{y_{i-1\,j-1} + y_{i-1\,j} + y_{ij}}{3}\right) = 0. \qquad \text{(IIID.2)}$$

When $N = 0$ (wave equation) this gives the explicit method

$$y_{i\,j+1} = \frac{h^2}{k^2}(y_{i+1\,j} - 2\,y_{ij} + y_{i-1\,j}) + 2\,y_{ij} - y_{i\,j-1}.$$

In MPS some numerical investigations of this integrator were undertaken using the sine-Gordon equation [(IIID.1) with $N'(\phi) = \sin\phi$] with periodic boundary conditions. Figure III.6 shows the time evolution of a waveform through a soliton collision just before the simulation stops, and may be compared to Figure III.7. As can be seen, the high frequency oscillations that are present during the soliton collisions are smaller and smoother for the triangle-based multisymplectic method than for the energy-conserving method of Guo et al. [1986]. ♦

The multisymplectic formalism appropriate for discrete elasticity was given in Marsden et al. [2001]. Motivated by this work, Lew et al. [2003] developed

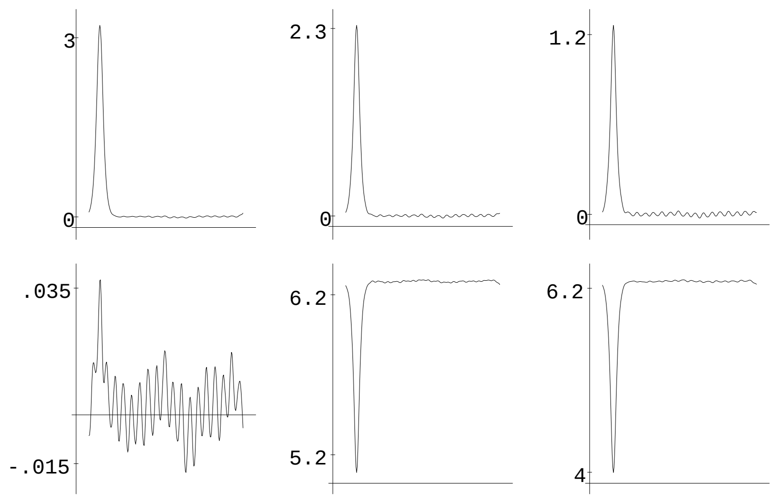


Figure III.6: A soliton collision after the triangle-based multisymplectic method (IIID.2) has simulated about 5000 soliton collisions. The solitons collide beginning at the top left and proceed to the top right, then to the bottom left, and finally to the bottom right.

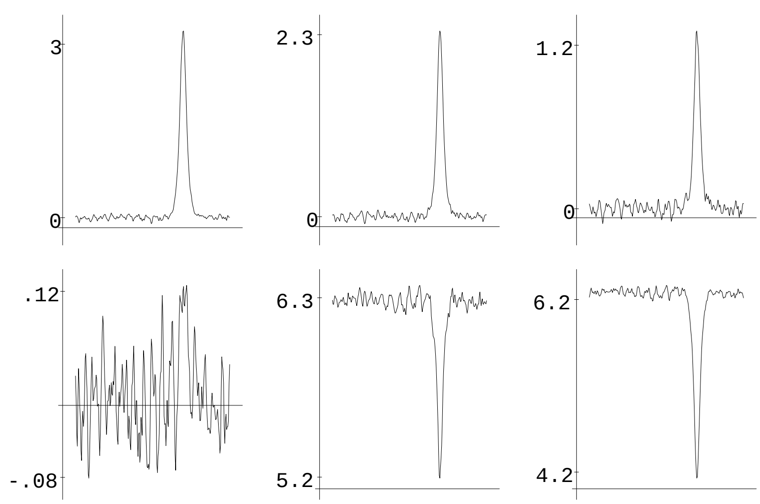


Figure III.7: Similar to the previous plot but using an energy-conserving integrator.

the theory of ***asynchronous variational integrators*** (***AVIs***) along with an implementation for the case of elastodynamics (discussed below). These integrators are based on the introduction of spacetime discretizations allowing different time steps for different elements in a finite element mesh. This is similar in spirit to subcycling (see, for example, Neal and Belytschko [1989]), but with no constraints on the ratio of time step between adjacent elements. It is the flexibility of the variational formulation (rather than trying to directly preserve the multisymplectic structure) that allows for such a formulation, without any sacrifice of the benefits of structured integrators.

A local discrete energy balance equation is obtained in a natural way in the AVI formalism. This equation can be satisfied exactly by adjusting the elemental time steps. However, it is sometimes computationally expensive to do so and from simulations (such as the one given below), it seems to be unnecessary. That is, the phenomenon of near energy conservation indefinitely in time appears to hold, just as in the finite-dimensional case. However, the full theory of a backward error analysis in the PDE context is in its infancy, cf. Moore and Reich [2003].

Example: Elastodynamics Simulation. The formulation and implementation of a sample algorithm (explicit Newmark for the time steps) has been carried out in this framework. An important and nontrivial issue is how to decide which elements to update next consistent with hyperbolicity (causality) and the CFL (Courant–Friedrichs–Levy) condition; one accomplishes this using the notion of a *priority queue* borrowed from computer science. Figure III.8 shows one snapshot of the dynamics of an elastic L-beam (the beam is undergoing oscillatory deformations). The smaller elements near the edges are updated much more frequently than the larger elements.

The figure also shows the very favorable energy behavior for the L-beam obtained with AVI techniques; the figure shows the total energy, but it is important to note that also the local energy balance is excellent—that is, there is no spurious energy exchange between elements as can be obtained with other elements. ⧫

Issues of small elements (sliver elements) are even more pronounced in other examples such as rotating elastic helicopter blades (without the hydrodynamics) which have also been simulated in some detail. The helicopter blade is one of the examples that was considered by the late Juan Simo who showed that standard

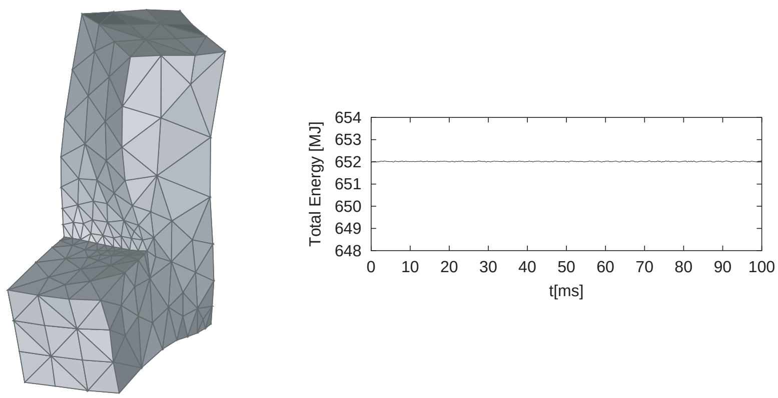


Figure III.8: AVI methods are used to simulate the dynamics of an elastic L-beam. The energy of the L-beam is nearly constant after a long integration run with millions of updates of the smallest elements.

(and even highly touted) algorithms can lead to troubles of various sorts. For example, if the modeling is not done carefully, then it can lead to spurious softening and also, even though the algorithm may be energy-respecting, it can be very bad as far as angular momentum conservation is concerned. The present AVI techniques suffer from none of these difficulties.

Amongst the many other possible future directions, one that is very exciting involves combining AVIs with emerging theories of *Discrete Exterior Calculus* (DEC) or *Discrete Tensor Calculus* (DTC). For example, it is known in computational electromagnetism (see, for instance Bossavit [1998]) that one gets spurious modes if the usual grad–div–curl relation is violated on the discrete level.

References

Abraham, R. and J. Marsden [1978], *Foundations of Mechanics*. Addison-Wesley, Menlo Park, California, second edition.

Abraham, R., J. Marsden, and T. Ratiu [1988], *Manifolds, Tensor Analysis, and Applications*. Springer-Verlag, New York, second edition.

Aldaya, V. and J. A. de Azcárraga [1980a], Geometric formulation of classical mechanics and field theory, *Rivista del Nuovo Cimento* **3**, 1–66.

Aldaya, V. and J. A. de Azcárraga [1980b], Higher-order Hamiltonian formalism in field theory, *J. Phys. A: Math. Gen.* **13**, 2545–2551.

Anco, S. C. and R. S. Tung [2002], Covariant Hamiltonian boundary conditions in General Relativity for spatially bounded space-time regions, *J. Math. Phys.* **43**, 5531–5566.

Anderson, I. [1992], Introduction to the variational bicomplex, *Contemp. Math.* **132**, 51–73.

Anderson, I. and J. M. Arms [1986], Perturbations of conservation laws in general relativity, *Ann. Phys.* **167**, 354–389.

Anderson, J. L. [1967], *Principles of Relativity Physics*. Academic Press, New York.

Andersson, L. [1987], Momenta and reduction for general relativity, *J. Geom. Phys.* **4**, 289–314.

Andersson, L. and V. Moncrief [2002], Elliptic-hyperbolic systems and the Einstein equations, *arXiv:* gr-qc/0110111.

Arms, J. M. [1977], Linearization stability of the Einstein–Maxwell system, *J. Math. Phys.* **18**, 830–833.

Arms, J. M. [1981], The structure of the solution set for the Yang–Mills equations, *Math. Proc. Camb. Phil. Soc.* **90**, 361–372.

Arms, J. M., R. Cushman, and M. J. Gotay [1991], A universal reduction procedure for Hamiltonian group actions. In Ratiu, T., editor, *The Geometry of Hamiltonian Systems*, volume 22 of *Mathematical Sciences Research Institute Publ.*, pages 33–51. Springer-Verlag, New York.

Arms, J. M., M. J. Gotay, and G. Jennings [1990], Geometric and algebraic reduction for singular momentum mappings, *Adv. in Math.* **79**, 43–103.

Arms, J. M., M. J. Gotay, and D. C. Wilbour [1989], Zero levels of momentum mappings for cotangent actions, *Nuc. Phys. B* (Proc. Suppl.) **6**, 384–389.

Arms, J. M. and J. E. Marsden [1979], The absence of Killing fields is necessary for linearization stability of Einstein's equations, *Ind. Univ. Math. J.* **28**, 119–125.

Arms, J. M., J. E. Marsden, and V. Moncrief [1981], Symmetry and bifurcations of momentum mappings, *Commun. Math. Phys.* **78**, 455–478.

Arms, J. M., J. E. Marsden, and V. Moncrief [1982], The structure of the space of solutions of Einstein's equations: II. Several Killing Fields and the Einstein–Yang–Mills equations, *Ann. Phys.* **144**, 81–106.

Arnowitt, R., S. Deser, and C. W. Misner [1962], The dynamics of general relativity. In Witten, L., editor, *Gravitation, an Introduction to Current Research*, pages 227–265. Wiley, New York.

Ashtekar, A. [1986], New variables for classical and quantum gravity, *Phys. Rev. Lett.* **57**, 2244–2247.

Ashtekar, A. [1987], New Hamiltonian formulation of general relativity, *Phys. Rev. D* **36**, 1587–1602.

Ashtekar, A., L. Bombelli, and O. Reula [1991], The covariant phase space of asymptotically flat gravitational fields. In Francaviglia, M., editor, *Mechanics, Analysis and Geometry: 200 Years After Lagrange*, pages 417–450. North-Holland, Amsterdam.

Bao, D. [1984], A sufficient condition for the linearization stability of $N = 1$ supergravity: a preliminary report, *Ann. Phys.* **158**, 211–278.

Bao, D., Y. Choquet-Bruhat, J. Isenberg, and P. Yasskin [1985], The well-posedness of $(N = 1)$ classical supergravity, *J. Math. Phys.* **26**, 329–333.

Bao, D., J. Isenberg, and P. B. Yasskin [1985], The dynamics of the Einstein–Dirac system. I. A principal bundle formulation of the theory and its canonical analysis, *Ann. Phys.* **164**, 103–171.

Bao, D., J. Marsden, and R. Walton [1985], The Hamiltonian structure of general relativistic perfect fluids, *Commun. Math. Phys.* **99**, 319–345.

Batlle, C., J. Gomis, and J. M. Pons [1986], Hamiltonian and Lagrangian constraints of the bosonic string, *Phys. Rev. D* **34**, 2430–2432.

Beig, R. [1989], On the classical geometry of bosonic string dynamics, *Int. J. Theor. Phys.* **30**, 211–224.

Beig, R. and B. G. Schmidt [2003], Relativistic elasticity, *Class. Quantum Grav.* **20**, 889–904.

Belinfante, F. J. [1939], On the spin angular momentum of mesons, *Physica* **vi**, 887–898.

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.

Bernal, A. N. and M. Sánchez [2005], Smooth globally hyperbolic splittings and temporal functions, *arXiv:* gr-qc/0404084.

Binz, E., J. Śniatycki, and H. Fischer [1988], *The Geometry of Classical Fields.* North Holland, Amsterdam.

Bloch, A. M., P. Crouch, J. E. Marsden, and T. S. Ratiu [2002], The symmetric representation of the rigid body equations and their discretization, *Nonlinearity* **15**, 1309–1341.

Bobenko, A. I. and Y. B. Suris [1999a], Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top, *Commun. Math. Phys.* **204**, 147–188.

Bobenko, A. and Y. Suris [1999b], Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products, *Lett. Math. Phys.* **49**, 79–93.

A. Bossavit [1998], *Computational Electromagnetism.* Academic Press, Boston.

Bridges, T. J. [1997], Multi-symplectic structures and wave propagation, *Math. Proc. Camb. Phil. Soc.* **121**, 147–190.

Bridges, T. J. [2006], Canonical multi-symplectic structure on the total exterior algebra bundle, *Proc. Roy. Soc. London A* **462**, 1531–1551.

Bridges, T. J. and S. Reich [2001], Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity, *Phys. Lett. A* **284**, 184–193.

Cantrijn, F., L. A. Ibort, and M. de León [1999], On the geometry of multisymplectic manifolds, *J. Aust. Math. Soc. A* **66**, 303–330.

Carathéodory, C. [1935], *Variationsrechnung und Partielle Differentialgleichungen erster Ordnung.* Teubner, Leipzig. (Reprinted by Chelsea, New York, 1982).

Cartan, E. [1922], *Leçons sur les Invariants Intégraux.* Hermann, Paris.

Castrillon-Lopez, M. and J. E. Marsden [2003], Some remarks on Lagrangian and Poisson reduction for field theories, *J. Geom. Phys.* **48**, 52–83.

Cendra, H. and M. Etchechoury [2005], Desingularization of implicit analytic differential equations, *arXiv:* math.CA/0507131.

Cendra, H., J.E. Marsden, and T.S. Ratiu [2001], *Lagrangian reduction by stages*, volume 152 of *Memoirs.* Amer. Math. Soc., Providence.

Chernoff, P. R. and J. E. Marsden [1974], *Properties of infinite dimensional Hamiltonian systems*, volume 425 of *Lecture Notes in Math.* Springer, New York.

Choquet-Bruhat, Y. [1962], The Cauchy Problem. In Witten, L., editor, *Gravitation: an Introduction to Current Research*, pages 130–168. Wiley, New York.

Choquet-Bruhat, Y. [2004], Causal evolution for Einstein gravitation, *Lecture Notes in Pure and Appl. Math.* **233**, 129–144.

Choquet-Bruhat, Y., A. E. Fischer, and J. E. Marsden [1979], Maximal hypersurfaces and positivity of mass. In Ehlers, J., editor, *Isolated Gravitating Systems and General Relativity*, pages 322–395. Italian Physical Society.

Corson, E. M. [1953], *Introduction to Tensors, Spinors and Relativistic Wave Equations.* Chelsea, New York.

Crnković, Č. and E. Witten [1987], Covariant description of canonical formalism in geometrical theories. In Hawking, S. W. and W. Israel, editors, *Newton's Tercentenary Volume*, pages 666–684. Cambridge Univ. Press, Cambridge.

Davis, W. R. [1970], *Classical Fields, Particles, and the Theory of Relativity.* Gordon and Breach, New York.

De Donder, T. [1930], *Théorie Invariantive du Calcul des Variations.* Gauthier-Villars, Paris.

de León, M. and P. R. Rodrigues [1985], *Generalized Classical Mechanics and Field Theory.* North Holland, Amsterdam.

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

Dedecker, P. [1957], Calcul des variations et topologie algébrique, *Mem. Soc. Roy. Sc. Liège* **19**, 1–216.

Dedecker, P. [1977], On the generalization of symplectic geometry to multiple integrals in the calculus of variations, *Lecture Notes in Math.* **570**, 395–456.

Deligne, P. and D. S. Freed [1999], Classical field theory. In Deligne, P., P. Etingof, D. S. Freed, L. C. Jeffrey, D. Kazhdan, J. W. Morgan, D. R. Morrison, and E. Witten, editors, *Quantum Fields and Strings: A Course for Mathematicians*, volume 1, pages 137–225. Amer. Math. Soc., Providence.

Desbrun, M., E. Kanso, J.E. Marsden, A. Mackenzie, and Y. Tong [2006], A stochastic approach to variational algorithms for fluid mechanics, *in preparation.*

Dirac, P. A. M. [1950], Generalized Hamiltonian dynamics, *Can. J. Math.* **2**, 129–148.

Dirac, P. A. M. [1958], The theory of gravitation in Hamiltonian form, *Proc. Roy. Soc. London A* **246**, 333–343.

Dirac, P. A. M. [1964], *Lectures on Quantum Mechanics.* Academic Press, New York.

Dutt, S. K. and M. Dresden [1986], Pure gravity as a constrained second order system, *SUNY preprint* ITP-SB-86-32.

Ebin, D. G. and J. E. Marsden [1970], Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. Math.* **92**, 102–163.

Echeverria-Enríques, A., M. C. Muñoz-Lecanda, and N. Román-Roy [2000], Geometry of multisymplectic Hamiltonian first-order field theories, *J. Math. Phys.* **41**, 7402–7444.

Esposito, G., G. Gionti, and C. Stornaiolo [1995], Spacetime covariant form of Ashtekar's constraints, *Nuovo Cimento* **B110**, 1137–1152.

Ferraris, M. and M. Francaviglia [1983], On the global structure of Lagrangian and Hamiltonian formalisms in higher order calculus of variations. In Modugno, M., editor, *Proc. of the Meeting "Geometry and Physics"*, pages 43–70. Pitagora Editrice, Bologna.

Ferraris, M. and M. Francaviglia [1985], Energy-momentum tensors and stress tensors in geometric field theories, *J. Math. Phys.* **26**, 1243–1252.

Ferraris, M. and M. Francaviglia [1991], The Lagrangian approach to conserved quantities in general relativity. In Francaviglia, M., editor, *Mechanics, Analysis, and Geometry: 200 Years After Lagrange*, pages 451–488. North Holland, Amsterdam.

Fetecau, R., J. E. Marsden, M. Ortiz, and M. West [2003], Nonsmooth Lagrangian mechanics and variational collision integrators, *SIAM J. Dynamical Systems* **2**, 381–416.

Fischer, A. E. [1982], A unified approach to conservation laws in general relativity, gauge theories, and elementary particle physics, *Gen. Rel. Grav.* **14**, 683–689.

Fischer, A. E. [1985], Conservation laws in gauge field theories. In Rassias, G. and T. Rassias, editors, *Differential Geometry, Calculus of Variations, and their Applications*, pages 211–253. Marcel Dekker, New York.

Fischer, A. E. and J. E. Marsden [1973], Linearization stability of the Einstein equations, *Bull. Amer. Math. Soc.* **79**, 995–1001.

Fischer, A. E. and J. E. Marsden [1975], Linearization stability of nonlinear partial differential equations. In volume 27 of *Proc. Symp. Pure Math.*, pages 219–263. Amer. Math. Soc., Providence.

Fischer, A. E. and J. E. Marsden [1979a], Topics in the dynamics of general relativity. In Ehlers, J., editor, *Isolated Gravitating Systems in General Relativity*, pages 322–395. Italian Physical Society.

Fischer, A. E. and J. E. Marsden [1979b], The initial value problem and the dynamical formulation of general relativity. In Hawking, S. W. and W. Israel, editors, *General Relativity*, pages 138–211. Cambridge Univ. Press, Cambridge.

Fischer, A. E., J. E. Marsden, and V. Moncrief [1980], The structure of the space of solutions of Einstein's equations I: One Killing field, *Ann. Inst. H. Poincaré* **33**, 147–194.

Forger, M., C. Paufler, and H. Römer [2003a], The Poisson bracket for Poisson forms in multisymplectic field theory, *Rev. Math. Phys.* **15**, 705–744.

Forger, M., C. Paufler, and H. Römer [2003b], A general construction of Poisson brackets on exact multisymplectic manifolds, *Rep. Math. Phys.* **51**, 187–195.

Forger, M. and H. Römer [2004], Currents and the energy-momentum tensor in classical field theory: A fresh look at an old problem, *Ann. Phys.* **309**, 306–389.

Fritelli, S. and O. Reula [1996], First-order symmetric-hyperbolic Einstein equations with arbitrary fixed gauge, *Phys. Rev. Lett.* **76**, 4667–4670.

García, P. L. [1974], The Poincaré–Cartan invariant in the calculus of variations, *Symp. Math.* **14**, 219–246.

García, P. L. and J. Muñoz [1991], Higher order regular variational problems. In Donato, P., C. Duval, J. Elhadad, and G. M. Tuynman, editors, *Symplectic Geometry and Mathematical Physics*, volume 99 of *Progress in Math.*, pages 136–159. Birkhäuser, Boston.

Gawędzki, K. [1972], On the geometrization of the canonical formalism in the classical field theory, *Rep. Math. Phys.* **3**, 307–326.

Goldschmidt, H. and S. Sternberg [1973], The Hamilton–Cartan formalism in the calculus of variations, *Ann. Inst. Fourier* **23**, 203–267.

Gotay, M. J. [1979], Presymplectic Manifolds, Geometric Constraint Theory and the Dirac–Bergmann Theory of Constraints, *Thesis.* Univ. of Maryland, Technical Report 80–063.

Gotay, M. J. [1983], On the validity of Dirac's conjecture regarding first class secondary constraints, *J. Phys. A: Math. Gen.* **16**, L141–145.

Gotay, M. J. [1984], Poisson reduction and quantization for the $(n+1)$-photon, *J. Math. Phys.* **25**, 2154–2159.

Gotay, M. J. [1988], A multisymplectic approach to the KdV equation. In Bleuler, K. and M. Werner, editors, *Differential Geometric Methods in Theoretical Physics*, volume 250 of *NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences*, pages 295–305. Kluwer, Dordrecht.

Gotay, M. J. [1989], Reduction of homogeneous Yang–Mills fields, *J. Geom. Phys.* **6**, 349–365.

Gotay, M. J. [1991a], A multisymplectic framework for classical field theory and the calculus of variations I. Covariant Hamiltonian formalism. In Francaviglia, M., editor, *Mechanics, Analysis, and Geometry: 200 Years After Lagrange*, pages 203–235. North Holland, Amsterdam.

Gotay, M. J. [1991b], A multisymplectic framework for classical field theory and the calculus of variations II. Space + time decomposition, *Diff. Geom. Appl.* **1**, 375–390.

Gotay, M. J. [1991c], An exterior differential systems approach to the Cartan form. In Donato, P., C. Duval, J. Elhadad, and G. Tuynman, editors, *Symplectic Geometry and Mathematical Physics*, volume 99 of *Progress in Math.*, pages 160–188. Birkhäuser, Boston.

Gotay, M. J. and J. Hanson [2007], The canonical analysis of second-order Einstein gravity, *in preparation.*

Gotay, M. J., R. Lashof, J. Śniatycki, and A. Weinstein [1983], Closed forms on symplectic fibre bundles, *Comment. Math. Helvetici* **58**, 617–621.

Gotay, M. J. and J. E. Marsden [1992], Stress-energy-momentum tensors and the Belinfante–Rosenfeld formula, *Contemp. Math.* **132**, 367–391.

Gotay, M. J. and J. M. Nester [1979], Presymplectic geometry, gauge transformations and the Dirac theory of constraints, *Lecture Notes in Phys.* **94**, 272–279.

Gotay, M. J. and J. M. Nester [1980], Generalized constraint algorithm and special presymplectic manifolds, *Lecture Notes in Math.* **775**, 78–104.

Gotay, M. J., J. M. Nester, and G. Hinds [1978], Presymplectic manifolds and the Dirac–Bergmann theory of constraints, *J. Math. Phys.* **19**, 2388–2399.

Green, M. B., J. H. Schwarz, and E. Witten [1987], *Superstring Theory*, volume I: Introduction of *Progress in Math.* Cambridge Univ. Press, Cambridge.

Griffiths, P. A. [1983], *Exterior Differential Systems and the Calculus of Variations*, volume 25 of *Progress in Math.* Birkhäuser, Boston.

Grundlach, C. and J. M. Martín-García [2004], Symmetric hyperbolic form of systems of second-order evolution equations subject to constraints, *arXiv:* gr-qc/0402079.

Guillemin, V. and S. Sternberg [1977], *Geometric Asymptotics*, volume 14 of *Amer. Math. Soc. Surveys.* Amer. Math. Soc., Providence.

Guillemin, V. and S. Sternberg [1984], *Symplectic Techniques in Physics.* Cambridge Univ. Press, Cambridge.

Günther, C. [1987], The polysymplectic Hamiltonian formalism in field theory and the calculus of variations, *J. Diff. Geom.* **25**, 23–53.

Guo, B. Y., P. J. Pascual, M. J. Rodriguez, and L. Vasquez [1986], Numerical solution of the sine-Gordon equation, *Appl. Math. Comput.* **18**, 1–14.

Hairer, E. and C. Lubich [2000], Long-time energy conservation of numerical methods for oscillatory differential equations, *SIAM J. Num. Anal.* **38**, 414–441 (electronic).

Hairer, E., C. Lubich, and G. Wanner [2001], *Geometric Numerical Integration.* Springer, Berlin.

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

Hawking, S. W. and G. F. R. Ellis [1973], *The Large Scale Structure of Spacetime.* Cambridge Univ. Press, Cambridge.

Hermann, R. [1968], *Differential Geometry and the Calculus of Variations.* Academic Press, Second Edition. (Reprinted by Math. Sci. Press, Brookline, Massachusetts).

Hermann, R. [1975], *Gauge Fields and Cartan–Ehresmann Connections Part A*. Math. Sci. Press, Brookline, Massachusetts.

Holm, D. [1985], Hamiltonian formalism of general relativistic adiabatic fluids, *Physica* **17D**, 1–36.

Holst, M., L. Lindblom, R. Owen, H. P. Pfeiffer, M. A. Scheel, and L. E. Kidder [2004], Optimal constraint projection for hyperbolic evolution systems, *Phys. Rev. D* **70**, 084017, 17.

Horák, M. and I. Kolář [1983], On the higher order Poincaré–Cartan forms, *Cz. Math. J.* **33**, 467–475.

Horowitz, G. [1989], Exactly soluble diffeomorphism invariant theories, *Commun. Math. Phys.* **129**, 417–437.

Hughes, T., T. Kato, and J. Marsden [1977], Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity, *Arch. Rat. Mech. Anal.* **63**, 273–294.

Isenberg, J. and J. Marsden [1982], A slice theorem for the space of solutions of Einstein's equations, *Phys. Rep.* **89**, 179–222.

Isenberg, J. and J. Nester [1977], The effect of gravitational interaction on classical fields: a Hamilton–Dirac analysis, *Ann. Phys.* **107**, 56–81.

Isenberg, J. and J. Nester [1980], Canonical analysis of relativistic field theories. In Held, A., editor, *General Relativity and Gravitation, Vol. 1*, pages 23–97. Plenum Press, New York.

Islas, A. L. and C. M. Schober [2004], On the preservation of phase space structure under multisymplectic discretization, *J. Comput. Phys.* **197**, 585–609.

Jalnapurkar, S.M., M. Leok, J.E. Marsden and M. West [2006], Discrete Routh reduction, *J. Phys. A: Math. Gen.*, **39**, 5521–5544.

John, F. [1982], *Partial Differential Equations.* Applied Mathematical Sciences **1**. Springer-Verlag, New York, fourth edition.

Johnson, J. E. [1985], Markov-type Lie groups in $GL(n, \mathbb{R})$, *J. Math. Phys.* **26**, 252–257.

Kanatchikov, I. V. [1997], On field theoretic generalizations of a Poisson algebra, *Rep. Math. Phys.* **40**, 225–234.

Kanatchikov, I. V. [1998], Canonical structure of classical field theory in the polymomentum phase space, *Rep. Math. Phys.* **41**, 49–90.

Kane, C., J. E. Marsden, M. Ortiz, and M. West [2000], Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems, *Int. J. Num. Math. Eng.* **49**, 1295–1325.

Kastrup, H. A. [1983], Canonical theories of Lagrangian dynamical systems in physics, *Phys. Rep.* **101**, 1–167.

Kijowski, J. [1973], A finite-dimensional canonical formalism in the classical field theory, *Commun. Math. Phys.* **30**, 99–128.

Kijowski, J. [1974], Multiphase spaces and gauge in the calculus of variations, *Bull. Acad. Sc. Polon.* **22**, 1219–1225.

Kijowski, J. and W. Szczyrba [1975], Multisymplectic manifolds and the geometrical construction of the Poisson brackets in the classical field theory. In Souriau, J.-M., editor, *Géométrie Symplectique et Physique Mathématique*, pages 347–379. C.N.R.S., Paris.

Kijowski, J. and W. Szczyrba [1976], A canonical structure for classical field theories, *Commun. Math. Phys.* **46**, 183–206.

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

Klainerman, S. and F. Nicoló [1999], On local and global aspects of the Cauchy problem in general relativity, *Class. Quantum Grav.* **16**, R73–R157.

Kolář, I. [1984], A geometric version of the higher order Hamilton formalism in fibered manifolds, *J. Geom. Phys.* **1**, 127–137.

Kosmann-Schwarzbach, Y. [1981], Hamiltonian systems on fibered manifolds, *Lett. Math. Phys.* **5**, 229–237.

Kouranbaeva, S. and S. Shkoller [2000], A variational approach to second-order multisymplectic field theory, *J. Geom. Phys.* **35**, 333–366.

Krupka, D. [1971], Lagrange theory in fibered manifolds, *Rep. Math. Phys.* **2**, 121–133.

Krupka, D. [1987], Geometry of Lagrangean structures. 3., *Supp. Rend. Circ. Mat.* (Palermo) ser. II, **14**, 187–224.

Krupka, D. and O. Štěpánková [1983], On the Hamilton form in second order calculus of variations. In Modugno, M., editor, *Proc. of the Meeting "Geometry and Physics"*, pages 85–101. Pitagora Editrice, Bologna.

Kuchař, K. [1973], Canonical quantization of gravity. In Israel, W., editor, *Relativity, Astrophysics and Cosmology*, pages 237–288. Reidel, Dordrecht.

Kuchař, K. [1974], Geometrodynamics regained: a Lagrangian approach, *J. Math. Phys.* **15**, 708–715.

Kuchař, K. [1976], Geometry of hyperspace I–IV, *J. Math. Phys.* **17, 18**, 777–820, 1589–1597.

Kummer, M. [1981], On the construction of the reduced space of a Hamiltonian system with symmetry, *Indiana Univ. Math. J.* **30**, 281–291.

Künzle, H. P. [1986], Lagrangian formalism for adiabatic fluids on five-dimensional spacetime, *Can. J. Phys.* **64**, 185–189.

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

Künzle, H. P. and J. M. Nester [1984], Hamiltonian formulation of gravitating perfect fluids and the Newtonian limit, *J. Math. Phys.* **25**, 1009–1018.

Kupershmidt, B. A. [1980], Geometry of jet bundles and the structure of Lagrangian and Hamiltonian formalisms, *Lecture Notes in Math.* **775**, 162–218.

Lawson, J. K. [2000], A frame bundle generalization of multisymplectic field theories, *Rep. Math. Phys.* **45**, 183–205.

Lepage, T. H. J. [1936], Sur les champs géodésiques du calcul des variations, *Bull. Acad. Roy. Belg., Cl. Sci.* **22**, 716–729, 1036–1046.

Lepage, T. H. J. [1941], Sur les champs géodésique des intégrales multiples, *Bull. Acad. Roy. Belg., Cl. Sci.* **27**, 27–46.

Lepage, T. H. J. [1942], Champs stationnaires, champs géodésiques et formes intégrables, *Bull. Acad. Roy. Belg., Cl. Sci.* **28**, 73–92, 247–265.

Lew, A., J. E. Marsden, M. Ortiz, and M. West [2003], Asynchronous variational integrators, *Arch. Rational Mech. Anal.* **167**, 85–146.

Lew, A., J. E. Marsden, M. Ortiz, and M. West [2004], Variational time integrators, *Int. J. Num. Meth. Engrg.* **60**, 153–212.

Lusanna, L. [1991], The second Noether theorem as the basis of the theory of singular Lagrangians and Hamiltonian constraints, *Riv. Nuovo Cimento* **14**(3), 1–75.

Lusanna, L. [1993], The Shanmugadhasan canonical transformation, function groups and the extended second Noether theorem, *Int. J. Mod. Phys. A* **8**, 4193–4233.

Marsden, J. E. [1988], The Hamiltonian formulation of classical field theory, *Contemp. Math.* **71**, 221–235.

Marsden, J. E. [1992], *Lectures in Mechanics*, volume 174 of *London Math. Soc. Lecture Notes*. Cambridge Univ. Press, Cambridge.

Marsden, J. E. and T. J. R. Hughes [1983], *Mathematical Foundations of Elasticity*. Prentice-Hall, Redwood City, California (Reprinted by Dover, 1994).

Marsden, J. E., R. Montgomery, P. J. Morrison, and W. B. Thompson [1986], Covariant Poisson brackets for classical fields, *Ann. Phys.* **169**, 29–48.

Marsden, J. E., R. Montgomery, and T. Ratiu [1990], *Symmetry, Reduction, and Phases in Mechanics*, volume 436 of *Mem. Amer. Math. Soc.*.

Marsden, J. E., G. W. Patrick, and S. Shkoller [1998], Multisymplectic geometry, variational integrators and nonlinear PDEs, *Comm. Math. Phys.* **199**, 351–395.

Marsden, J. E., S. Pekarsky, and S. Shkoller [1999], Discrete Euler–Poincaré and Lie–Poisson equations, *Nonlinearity* **12**, 1647–1662.

Marsden, J. E., S. Pekarsky, and S. Shkoller [2000], Symmetry reduction of discrete Lagrangian mechanics on Lie groups, *J. Geom. and Phys.* **36**, 140–151.

Marsden, J. E., S. Pekarsky, S. Shkoller, and M. West [2001], Variational methods, multisymplectic geometry and continuum mechanics, *J. Geom. Phys.* **38**, 253–284.

Marsden, J. E. and T. S. Ratiu [1999], *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Math.*. Springe, New York, second edition.

Marsden, J. E. and S. Shkoller [1999], Multisymplectic geometry, covariant Hamiltonians and water waves, *Math. Proc. Camb. Phil. Soc.* **125**, 553–575.

Marsden, J. E., A. Weinstein, T. Ratiu, R. Schmid, and R. G. Spencer [1983], Hamiltonian systems with symmetry, coadjoint orbits and plasma physics. In Benenti, S., M. Francaviglia, and A. Lichnerowicz, editors, *Proc. IUTAM–ISIMM Symposium on Modern Developments in Analytical Mechanics*, volume 117 of *Atti della Accademia delle Scienze di Torino*, pages 289–340.

Marsden, J. E. and M. West [2001], Discrete mechanics and variational integrators, *Acta Numerica* **10**, 357–514.

Martin, G. [1988], A Darboux theorem for multi-symplectic manifolds, *Lett. Math. Phys.* **16**, 133–138.

Martínez, E. [2004], Classical field theory on Lie algebroids: Multisymplectic formalism, *arXiv:* math.DG/0411352.

McOwen, R. C. [2003], *Partial Differential Equations: Methods and Applications.* Prentice Hall, Upper Saddle River, New Jersey, second edition.

Mikami, K. and A. Weinstein [1988], Moments and reduction for symplectic groupoid actions, *Publ. RIMS Kyoto Univ.* **24**, 121–140.

Misner, C. W., K. Thorne, and J. A. Wheeler [1973], *Gravitation.* W.H. Freeman, San Francisco.

Moncrief, V. [1975], Spacetime symmetries and linearization stability of the Einstein equations. I, *J. Math. Phys.* **16**, 493–498.

Moncrief, V. [1976], Spacetime symmetries and linearization stability of the Einstein equations. II, *J. Math. Phys.* **17**, 1892–1902.

Moore, B. and S. Reich [2003], Backward error analysis for multi-symplectic integration methods, *Num. Math.* **95**, 625–652.

Moreau, J.-J. [1982], Fluid dynamics and the calculus of horizontal variations, *Int. J. Engrg. Sci.* **20**, 389–411.

Moser, J. and A. P. Veselov [1991], Discrete versions of some classical integrable systems and factorization of matrix polynomials, *Comm. Math. Phys.* **139**, 217–243.

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 & Structures*, **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*, Colloq. Math. Soc. J. Bolyai **31**, pages 503–509. North-Holland, Amsterdam.

Olver, P. J. [1993], *Applications of Lie Groups in Differential Equations*, volume 107 of *Graduate Texts in Math.* Springer, New York, second edition.

Ortega, J.-P. and T. S. Ratiu [2004], *Momentum Maps and Hamiltonian Reduction*, volume 222 of *Progress in Math.* Birkhäuser, Boston.

Ouzilou, R. [1972], Expression symplectique des problèmes variationnels, *Symp. Math.* **14**, 85–98.

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

Pâquet, P. V. [1941], Les formes differentiélles extérieurs Ω_n dans le calcul des variations, *Bull. Acad. Roy. Belg., Cl. Sci.* **27**, 65–84.

Patrick, G. W. [1985], Singular Momentum Mappings, Presymplectic Dynamics and Gauge Groups. *Thesis*, Univ. of Calgary.

Polyakov, A. M. [1981], Quantum geometry of bosonic strings, *Phys. Lett.* **103B**, 207–210.

Ragionieri, R. and R. Ricci [1981], Hamiltonian formalism in the calculus of variations, *Bollettino U.M.I.* **18-B**, 119–130.

Regge, T. and C. Teitelboim [1974], Role of surface integrals in the Hamiltonian formulation of general relativity, *Ann. Phys.* **88**, 286–318.

Rosenfeld, L. [1940], Sur le tenseur d'impulsion-énergie, *Mém. Acad. Roy. Belg. Sci.* **18**, 1–30.

Sardanashvily, G. [1993], *Gauge Theory in Jet Manifolds.* Hadronic Press, Palm Harbor, Florida.

Saunders, D. J. [1989], *The Geometry of Jet Bundles*, volume 142 of *London Math. Soc. Lecture Note Series.* Cambridge Univ. Press, Cambridge.

Saunders, D. J. [1992], The regularity of variational problems, *Contemp. Math.* **132**, 573–593.

Scherk, J. [1975], An introduction to the theory of dual models and strings, *Rev. Mod. Phys.* **47**, 123–164.

Shadwick, W. F. [1982], The Hamiltonian formulation of regular rth-order Lagrangian field theories, *Lett. Math. Phys.* **6**, 409–416.

Simo, J. C., D. Lewis, and J. Marsden [1991], Stability of relative equilibria I. The reduced energy momentum method, *Arch. Rat. Mech. Anal.* **115**.

Simo, J. C. and J. E. Marsden [1984], On the rotated stress tensor and a material version of the Doyle–Ericksen formula, *Arch. Rat. Mech. Anal.* **86**, 213–231.

Sjamaar, R. and E. Lerman [1991], Stratified symplectic spaces and reduction, *Ann. Math.* **134**, 375–422.

Śniatycki, J. [1970a], On the geometric structure of classical field theory in Lagrangian formulation, *Proc. Camb. Phil. Soc.* **68**, 475–484.

Śniatycki, J. [1970b], On the canonical formulation of general relativity. In *Proc. Journées Relativistes 1970*, pages 127–135. Faculté des Sciences, Caen.

Śniatycki, J. [1974], Dirac brackets in geometric dynamics, *Ann. Inst. H. Poincaré* **A20**, 365–372.

Śniatycki, J. [1984], The Cauchy data space formulation of classical field theory, *Rep. Math. Phys.* **19**, 407–422.

Śniatycki, J. [1988], Conservation laws in asymptotically flat spacetimes revisted, *Rep. Math. Phys.* **25**, 127–140.

Śniatycki, J. [2004], Multisymplectic reduction for proper actions, *Can. J. Math.* **56**, 638–654.

Soper, D. E. [1976], *Classical Field Theory.* Wiley-Interscience, New York.

Souriau, J.-M. [1974], Modèle de particule à spin dans le champ éléctromagnétique et gravitationnel, *Ann. Inst. H. Poincaré* **20**, 315–364.

Sternberg, S. [1977], Some preliminary remarks on the formal variational calculus of Gel'fand and Dikki, *Lecture Notes in Math.* **676**, 399–407.

Sundermeyer, K. [1982], Constrained dynamics, *Lecture Notes in Phys.* **169**.

Suris, Y. B [1989], The canonicity of mappings generated by Runge–Kutta type methods when integrating the system $\ddot{x} = -\partial u/\partial x$, *USSR Comput. Math. Phys.* **29**, 138–144.

Suris, Y. B. [1990], Hamiltonian methods of Runge-Kutta type and their variational interpretation, *Mat. Model.* **2**, 78–87.

Suris, Y. B. [2003], *The Problem of Integrable Discretization: Hamiltonian Approach.* volume 219 of *Progress in Math.* Birkhäuser, Boston.

Szczyrba, W. [1976a], Lagrangian formalism in the classical field theory, *Ann. Polon. Math.* **32**, 145–185.

Szczyrba, W. [1976b], A symplectic structure of the set of Einstein metrics: a canonical formalism for general relativity, *Commun. Math. Phys.* **51**, 163–182.

Thoutireddy, P. and M. Ortiz [2004], A variational r-adaption and shape-optimization method for finite-deformation elasticity, *Internat. J. Num. Methods Engrg.* **61**, 1–21.

Trautman, A. [1965], Foundations and current problems of general relativity. In *Lectures on General Relativity, Vol. I.* Prentice-Hall, Englewood Cliffs, New Jersey.

Trautman, A. [1967], Noether equations and conservation laws, *Commun. Math. Phys.* **6**, 248–261.

Wald, R. M. [1984], *General Relativity.* Univ. of Chicago Press, Chicago.

Wendlandt, J. M. and J. E. Marsden [1997], Mechanical integrators derived from a discrete variational principle, *Physica D* **106**, 223–246.

Wentzel, G. [1949], *Quantum Theory of Fields.* Interscience, New York.

Weyl, H. [1935], Geodesic fields in the calculus of variation for multiple integrals, *Ann. Math.* **36**, 607–629.

Yavari, A., J. E. Marsden, and M. Ortiz [2006], On spatial and material covariant balance laws in elasticity, *J. Math. Phys.* **47**, 1–53.

Yoshimura, H. and J. E. Marsden [2006], Dirac Structures and Lagrangian Mechanics, *J. Geom. Phys.*, *published online*, April, 2006.

Zuckerman, G. J. [1987], Action principles and global geometry. In Yau, S. T., editor, *Mathematical Aspects of String Theory*, volume 1 of *Adv. Ser. Math. Phys.*, pages 259–284. World Scientific, Singapore.