Momentum Maps and Classical Fields

Part II: Canonical Analysis of Field Theories

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

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

With the collaboration of

James Isenberg **Richard Montgomery**

With the scientific input of

Jędrzej Śniatycki **Philip B.Yasskin**

July 19, 2006

arXiv: math-ph/0411036

Contents

*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.

II—Canonical Analysis of Field Theories

With the covariant formulation in hand from the first part of this book, we begin in this second part to study the canonical (or "instantaneous") formulation of classical field theories. The canonical formulation works with fields defined as time-evolving cross sections of bundles over a Cauchy surface, rather than as sections of bundles over spacetime as in the covariant formulation. More precisely, for a given classical field theory, the (infinite-dimensional) instantaneous configuration space consists of the set $\mathcal{Y}_\Sigma$ of all smooth sections of a specified bundle Y_Σ over a Cauchy surface Σ, and a solution to the field equations is represented by a trajectory in $\mathcal{Y}_\Sigma$. As in classical mechanics, the Lagrangian formulation of the field equations of a classical field theory is defined on the tangent bundle $T\mathcal{Y}_\Sigma$, and the Hamiltonian formulation is defined on the cotangent bundle $T^*\mathcal{Y}_\Sigma$, which has a canonically defined symplectic structure ω_Σ.

To relate the canonical and the covariant approaches to classical field theory, we start in Chapter 5 by discussing embeddings $\Sigma \to X$ of Cauchy surfaces in spacetime, and considering the corresponding pull-back bundles $Y_\Sigma \to \Sigma$ of the covariant configuration bundle $Y \to X$. We go on in the same chapter to relate the covariant multisymplectic geometry of (Z, Ω) to the instantaneous symplectic geometry of $(T^*\mathcal{Y}_\Sigma, \omega_\Sigma)$ by showing that the multisymplectic form Ω on Z naturally induces the symplectic form ω_Σ on $T^*\mathcal{Y}_\Sigma$.

The discussion in Chapter 5 concerns primarily kinematical structures, such as spaces of fields and their geometries, but does not involve the action principle or the field equations for a given classical field theory. In Chapter 6, we proceed to consider field dynamics. A crucial feature of our discussion here is the degeneracy of the Lagrangian functionals for the field theories of interest. As a consequence of this degeneracy, we have constraints on the choice of initial data, and gauge freedom in the evolution of the fields. Chapter 6 considers the role of initial value constraints and gauge transformations in field dynamics. The discussion is framed primarily in the Hamiltonian formulation of the dynamics.

One of the primary goals of this work is to show how momentum maps are used in classical field theories which have both initial value constraints and gauge freedom. In Chapter 7, we begin to do this by describing how the covariant momentum maps defined on the multiphase space Z in Part I induce a generalization of momentum maps—"energy-momentum maps"—on the instantaneous phase spaces $T^*\mathcal{Y}_\Sigma$. We show that for a group action which leaves the Cauchy surface invariant, this energy-momentum map coincides with the usual notion of a momentum map. We also show, when the gauge group "includes" the spacetime diffeomorphism group, that one of the components of the energy-momentum map corresponding to spacetime diffeomorphisms can be identified (up to sign) with the Hamiltonian for the theory.

5 Symplectic Structures Associated with Cauchy Surfaces

The transition from the covariant to the instantaneous formalism once a Cauchy surface (or a foliation by Cauchy surfaces) has been chosen is a central ingredient of this work. It will eventually be used to cast the field dynamics into adjoint form and to determine when the first class constraint set (in the sense of Dirac) is the zero set of an appropriate energy-momentum map.

5A Cauchy Surfaces and Spaces of Fields

In any particular field theory, we assume there is singled out a class of hypersurfaces which we call ***Cauchy surfaces***. We will not give a precise definition here, but our usage of the term is intended to correspond to its meaning in general relativity (see, for instance, Hawking and Ellis [1973]).

Let Σ be a compact (oriented, connected) boundaryless n-manifold. We denote by $\mathrm{Emb}(\Sigma, X)$ the space of all smooth embeddings of Σ into X. (If the $(n+1)$-dimensional "spacetime" X carries a nonvariational Lorentz metric, we then understand $\mathrm{Emb}(\Sigma, X)$ to be the space of smooth *spacelike* embeddings of Σ into X.) As usual, many of the formal aspects of the constructions also work in the noncompact context with asymptotic conditions appropriate to the allowance of the necessary integrations by parts. However, the analysis necessary to cover the noncompact case need not be trivial; these considerations are important when dealing with isolated systems or asymptotically flat spacetimes. See Regge and Teitelboim [1974], Choquet-Bruhat et al. [1979], Śniatycki [1988], and Ashtekar et al. [1991].

For $\tau \in \mathrm{Emb}(\Sigma, X)$, let $\Sigma_\tau = \tau(\Sigma)$. The hypersurface Σ_τ will eventually be a Cauchy surface for the dynamics; we view Σ as a ***reference*** or ***model***

Cauchy surface. We will not need to topologize $\mathrm{Emb}(\Sigma, X)$ in this book; however, we note that when completed in appropriate C^k or Sobolev topologies, $\mathrm{Emb}(\Sigma, X)$ and other manifolds of maps introduced below are known to be smooth manifolds (see, for example, Palais [1968] and Ebin and Marsden [1970]).

If $\pi_{XK} : K \to X$ is a fiber bundle over X, then the space of smooth sections of the bundle will be denoted by the corresponding script letter, in this case $\mathcal{K}$. Occasionally, when this notation might be confusing, we will resort to the notation $\Gamma(K)$ or $\Gamma(X, K)$. We let K_τ denote the restriction of the bundle K to $\Sigma_\tau \subset X$ and let the corresponding script letter denote the space of its smooth sections, in this case $\mathcal{K}_\tau$. The collection of all $\mathcal{K}_\tau$ as τ ranges over $\mathrm{Emb}(\Sigma, X)$ forms a bundle over $\mathrm{Emb}(\Sigma, X)$ which we will denote $\mathcal{K}^\Sigma$.

The ***tangent space*** to $\mathcal{K}$ at a point σ is given by

$$T_\sigma \mathcal{K} = \left\{ W : X \to VK \;\middle|\; W \text{ covers } \sigma \right\}, \tag{5A.1}$$

where VK denotes the vertical tangent bundle of K. See Figure 5.1.

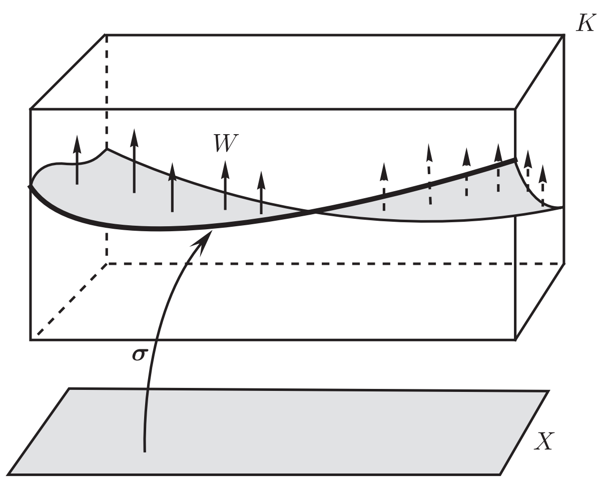

Figure 5.1: A tangent vector $W \in T_\sigma \mathcal{K}$

Similarly, the ***smooth cotangent space*** to $\mathcal{K}$ at σ is

$$T^*_\sigma \mathcal{K} = \left\{ \pi : X \to L(VK, \Lambda^{n+1} X) \;\middle|\; \pi \text{ covers } \sigma \right\}, \tag{5A.2}$$

where $L(VK, \Lambda^{n+1} X)$ is the vector bundle over K whose fiber at $k \in K_x$ is the set of linear maps from $V_k K$ to $\Lambda^{n+1}_x X$. The natural pairing of $T^*_\sigma \mathcal{K}$ with $T_\sigma \mathcal{K}$

is given by integration:

$$\langle \pi, V \rangle = \int_X \pi(V). \tag{5A.3}$$

One obtains similar formulas for $\mathcal{K}_\tau$ from the above by replacing X with Σ_τ and K with K_τ throughout (and replacing $n+1$ by n in (5A.2)). See Figure 5.2.

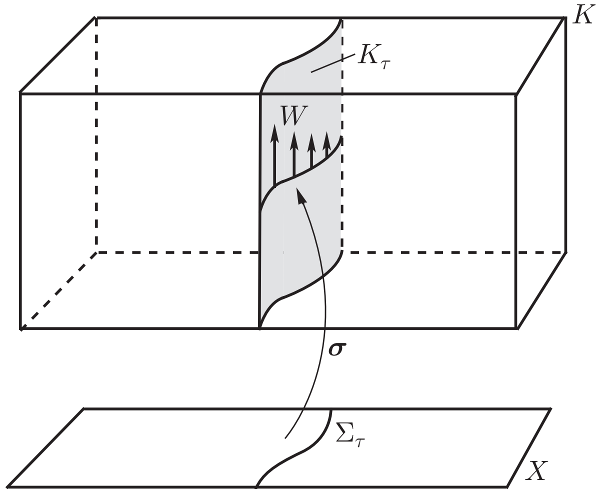

Figure 5.2: A tangent vector $W \in T_\sigma \mathcal{K}_\tau$

If ξ_K is any π_{XK}-projectable vector field on K, we define the ***Lie derivative*** of $\sigma \in \mathcal{K}$ along ξ_K to be the element of $T_\sigma \mathcal{K}$ given by

$$\pounds_{\xi_K} \sigma = T\sigma \circ \xi_X - \xi_K \circ \sigma. \tag{5A.4}$$

Note that $-\pounds_{\xi_K}\sigma$ is exactly the vertical component of $\xi_K \circ \sigma$. In coordinates (x^μ, k^A) on K we have

$$(\pounds_{\xi_K}\sigma)^A = \sigma^A{}_{,\mu}\xi^\mu - \xi^A \circ \sigma, \tag{5A.5}$$

where $\xi_K = (\xi^\mu, \xi^A)$.

Finally, if f is a map $\mathcal{K} \to \mathcal{F}(X)$ we define the "formal" partial derivatives $D_\mu f : \mathcal{K} \to \mathcal{F}(X)$ via

$$D_\mu f(\sigma) = f(\sigma)_{,\mu}. \tag{5A.6}$$

Intrinsically, this is the coordinate representation of the differential of the real valued function $f(\sigma)$.

5B Canonical Forms on $T^*\mathcal{Y}_\tau$ and $\mathcal{Z}_\tau$

In the instantaneous formalism the configuration space at "time" $\tau \in \mathrm{Emb}(\Sigma, X)$ will be denoted $\mathcal{Y}_\tau$, hereafter called the τ-***configuration space***. Likewise, the τ-***phase space*** is $T^*\mathcal{Y}_\tau$, the smooth cotangent bundle of $\mathcal{Y}_\tau$ with its canonical one-form θ_τ and canonical two-form ω_τ. These forms are defined using the same construction as for ordinary cotangent bundles (see Abraham and Marsden [1978] or Chernoff and Marsden [1974]). Specifically, we define θ_τ by

$$\theta_\tau(\varphi, \pi)(V) = \int_{\Sigma_\tau} \pi(T\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} \cdot V) \tag{5B.1}$$

where (φ, π) denotes a point in $T^*\mathcal{Y}_\tau$, $V \in T_{(\varphi,\pi)}T^*\mathcal{Y}_\tau$ and $\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} : T^*\mathcal{Y}_\tau \to \mathcal{Y}_\tau$ is the cotangent bundle projection. We define

$$\omega_\tau = -\mathbf{d}\theta_\tau. \tag{5B.2}$$

We now develop coordinate expressions for these forms. To this end choose a chart$\left(x^0, x^1, \ldots, x^n\right)$ on X which is ***adapted*** to τ in the sense that Σ_τ is locally a level set of x^0. Then an element $\pi \in T^*_\varphi\mathcal{Y}_\tau$, regarded as a map $\pi : \Sigma_\tau \to L(VY_\tau, \Lambda^n\Sigma_\tau)$, is expressible as

$$\pi = \pi_A \, dy^A \otimes d^n x_0, \tag{5B.3}$$

so for the canonical one- and two-forms on $T^*\mathcal{Y}_\tau$ we get

$$\theta_\tau(\varphi, \pi) = \int_{\Sigma_\tau} \pi_A \, d\varphi^A \otimes d^n x_0 \tag{5B.4}$$

and

$$\omega_\tau(\varphi, \pi) = \int_{\Sigma_\tau} (d\varphi^A \wedge d\pi_A) \otimes d^n x_0. \tag{5B.5}$$

For example, if $V \in T_{(\varphi,\pi)}(T^*\mathcal{Y}_\tau)$ is given in adapted coordinates by $V = (V^A, W_A)$, then we have

$$\theta_\tau(\varphi, \pi)(V) = \int_{\Sigma_\tau} \pi_A V^A d^n x_0.$$

To relate the symplectic manifold $T^*\mathcal{Y}_\tau$ to the multisymplectic manifold Z, we first use the multisymplectic structure on Z to induce a presymplectic structure on $\mathcal{Z}_\tau$ and then identify $T^*\mathcal{Y}_\tau$ with the quotient of $\mathcal{Z}_\tau$ by the kernel of this presymplectic form. Specifically, define the ***canonical one-form*** Θ_τ on $\mathcal{Z}_\tau$ by

$$\Theta_\tau(\sigma)(V) = \int_{\Sigma_\tau} \sigma^*(\mathbf{i}_V \Theta), \tag{5B.6}$$

where $\sigma \in \mathcal{Z}_\tau$, $V \in T_\sigma \mathcal{Z}_\tau$, and Θ is the canonical $(n+1)$-form on Z given by (2B.9). The ***canonical two-form*** Ω_τ on $\mathcal{Z}_\tau$ is

$$\Omega_\tau = -\mathbf{d}\Theta_\tau. \tag{5B.7}$$

Lemma 5B.1. *At $\sigma \in \mathcal{Z}_\tau$ and with Ω given by* (2B.10)*, we have*

$$\Omega_\tau(\sigma)(V, W) = \int_{\Sigma_\tau} \sigma^*(\mathbf{i}_W \mathbf{i}_V \Omega). \tag{5B.8}$$

Proof. Extend V, W to vector fields $\mathcal{V}, \mathcal{W}$ on $\mathcal{Z}_\tau$ by fixing π_{XZ}-vertical vector fields v, w on Z_τ such that $V = v \circ \sigma$ and $W = w \circ \sigma$ and letting $\mathcal{V}(\rho) = v \circ \rho$ and $\mathcal{W}(\rho) = w \circ \rho$ for $\rho \in \mathcal{Z}_\tau$. Note that if f_λ is the flow of w, $\mathcal{F}_\lambda(\rho) = f_\lambda \circ \rho$ is the flow of $\mathcal{W}$. Then, from the definition of the bracket in terms of flows, one finds that

$$[\mathcal{V}, \mathcal{W}](\rho) = [v, w] \circ \rho.$$

The derivative of $\Theta_\tau(\mathcal{V})$ along $\mathcal{W}$ at σ is

$$\begin{aligned}\mathcal{W}\left[\Theta_\tau(\mathcal{V})\right](\sigma) &= \frac{d}{d\lambda}\left[\Theta_\tau(\mathcal{V}) \circ \mathcal{F}_\lambda(\sigma)\right]\Big|_{\lambda=0} = \frac{d}{d\lambda}\left[\int_{\Sigma_\tau} \mathcal{F}_\lambda(\sigma)^*(\mathbf{i}_v\Theta)\right]\Big|_{\lambda=0} \\ &= \frac{d}{d\lambda}\left[\int_{\Sigma_\tau} \sigma^* f_\lambda^*(\mathbf{i}_v\Theta)\right]\Big|_{\lambda=0} = \int_{\Sigma_\tau} \sigma^*[\pounds_w \mathbf{i}_v \Theta].\end{aligned}$$

Thus, at $\sigma \in \mathcal{Z}_\tau$,

$$\begin{aligned}\mathbf{d}\Theta_\tau(\mathcal{V}, \mathcal{W}) &= \mathcal{V}\left[\Theta_\tau(\mathcal{W})\right] - \mathcal{W}[\Theta_\tau(\mathcal{V})] - \Theta_\tau([\mathcal{V}, \mathcal{W}]) \\ &= \int_{\Sigma_\tau} \sigma^*\left[\pounds_v \mathbf{i}_w \Theta - \pounds_w \mathbf{i}_v \Theta - \mathbf{i}_{[v,w]}\Theta\right] \\ &= \int_{\Sigma_\tau} \sigma^*(-\mathbf{d}\mathbf{i}_w \mathbf{i}_v \Theta + \mathbf{i}_w \mathbf{i}_v \mathbf{d}\Theta),\end{aligned}$$

and the first term vanishes by the definitions of Z and Θ, as both v, w are π_{XZ}-vertical.[11] ■

The two-form Ω_τ on $\mathcal{Z}_\tau$ is closed, but it has a nontrivial kernel, as the following development will show.

[11] This term also vanishes by Stokes' theorem, but in fact (5B.8) holds regardless of whether Σ_τ is compact and boundaryless.

5C Reduction of $\mathcal{Z}_\tau$ to $T^*\mathcal{Y}_\tau$

Our next goal is to prove that $\mathcal{Z}_\tau / \ker \Omega_\tau$ is canonically isomorphic to $T^*\mathcal{Y}_\tau$ and that the inherited symplectic form on the former is isomorphic to the canonical one on the latter. To do this, define a vector bundle map $R_\tau : \mathcal{Z}_\tau \to T^*\mathcal{Y}_\tau$ over $\mathcal{Y}_\tau$ by

$$\langle R_\tau(\sigma), V\rangle = \int_{\Sigma_\tau} \varphi^*(\mathbf{i}_V \sigma), \tag{5C.1}$$

where $\varphi = \pi_{YZ} \circ \sigma$ and $V \in T_\varphi \mathcal{Y}_\tau$; the integrand in (5C.1) at a point $x \in \Sigma_\tau$ is the interior product of $V(x)$ with $\sigma(x)$, resulting in an n-form on Y, which is then pulled back along φ to an n-form on Σ_τ at x. Interpreted as a map of Σ_τ to $L(VY_\tau, \Lambda^n \Sigma_\tau)$ which covers φ, $R_\tau(\sigma)$ is given by

$$\langle R_\tau(\sigma)(x), v\rangle = \varphi^* \mathbf{i}_v \sigma(x), \tag{5C.2}$$

where $v \in V_{\varphi(x)} Y_\tau$. In adapted coordinates, $\sigma \in \mathcal{Z}_\tau$ takes the form

$$(p_A{}^\mu \circ \sigma)\, dy^A \wedge d^n x_\mu + (p \circ \sigma)\, d^{n+1}x, \tag{5C.3}$$

and so we may write

$$R_\tau(\sigma) = (p_A{}^0 \circ \sigma)\, dy^A \otimes d^n x_0. \tag{5C.4}$$

Comparing (5C.4) with (5B.3), we see that the ***instantaneous momenta*** π_A correspond to the temporal components of the multimomenta $p_A{}^\mu$. Moreover, R_τ is obviously a surjective submersion with

$$\ker R_\tau = \left\{ \sigma \in \mathcal{Z}_\tau \mid p_A{}^0 \circ \sigma = 0 \right\}.$$

Remark 5C.1. Although we have defined R_τ as a map on sections from $\mathcal{Z}_\tau$ to $T^*\mathcal{Y}_\tau$, in actuality R_τ is a pointwise operation. We may in fact write (5C.2) as $R_\tau(\sigma) = r_\tau \circ \sigma$, where

$$r_\tau : Z_\tau \to V^*Y_\tau \otimes \Lambda^n \Sigma_\tau$$

is a bundle map over Y_τ. From (5C.3) and (5C.4), we see that in coordinate form $r_\tau(p, p_A{}^\mu) = p_A{}^0$ with

$$\ker r_\tau = \left\{ p_A{}^i dy^A \otimes d^n x_i + p\, d^{n+1}x \in Z_\tau \right\}. \qquad \blacklozenge$$

Proposition 5C.2. *We have*

$$R_\tau^* \theta_\tau = \Theta_\tau. \tag{5C.5}$$

Proof. Let $V \in T_\sigma \mathcal{Z}_\tau$. By the definitions of pull-back and the canonical one-form,

$$\langle (R_\tau^* \theta_\tau)(\sigma), V\rangle = \langle \theta_\tau(R_\tau(\sigma)), TR_\tau \cdot V\rangle = \langle R_\tau(\sigma), T\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} \cdot TR_\tau \cdot V\rangle.$$

However, since R_τ covers the identity,

$$\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} \circ R_\tau = \pi_{\mathcal{Y}_\tau, \mathcal{Z}_\tau}$$

and so

$$T\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} \cdot TR_\tau \cdot V = T\pi_{\mathcal{Y}_\tau, \mathcal{Z}_\tau} \cdot V = T\pi_{YZ} \circ V.$$

Thus by (5C.1), with $\varphi = \pi_{YZ} \circ \sigma$,

$$\begin{aligned}\langle R_\tau^* \theta_\tau(\sigma), V\rangle = \langle R_\tau(\sigma), T\pi_{YZ} \circ V\rangle &= \int_{\Sigma_\tau} \varphi^*((T\pi_{YZ} \circ V) \,\lrcorner\, \sigma)\\ &= \int_{\Sigma_\tau} \sigma^* \pi_{YZ}^*((T\pi_{YZ} \circ V) \,\lrcorner\, \sigma)\\ &= \int_{\Sigma_\tau} \sigma^*(V \,\lrcorner\, \pi_{YZ}^* \sigma).\end{aligned}$$

However, by (2B.7) and (2B.9), $\pi_{YZ}^*\sigma = \Theta \circ \sigma$. Thus by (5B.6),

$$\langle R_\tau^* \theta_\tau(\sigma), V\rangle = \langle \Theta_\tau(\sigma), V\rangle\,. \qquad \blacksquare$$

Corollary 5C.3.

(i) $R_\tau^* \omega_\tau = \Omega_\tau$.

(ii) $\ker T_\sigma R_\tau = \ker \Omega_\tau(\sigma)$.

(iii) *The induced quotient map $\mathcal{Z}_\tau / \ker R_\tau = \mathcal{Z}_\tau / \ker \Omega_\tau \to T^*\mathcal{Y}_\tau$ is a symplectic diffeomorphism.*

Proof. (i) follows by taking the exterior derivative of (5C.5). (ii) follows from (i), the (weak) nondegeneracy of ω_τ, the definition of pull-back and the fact that R_τ is a submersion. Finally, (iii) follows from (i), (ii), and the fact that R_τ is a surjective vector bundle map between vector bundles over $\mathcal{Y}_\tau$. $\blacksquare$

Thus, for each Cauchy surface Σ_τ, the multisymplectic structure Ω on Z induces a presymplectic structure Ω_τ on $\mathcal{Z}_\tau$, and this in turn induces the canonical symplectic structure ω_τ on the instantaneous phase space $T^*\mathcal{Y}_\tau$. Alternative constructions of Θ_τ and ω_τ are given in Zuckerman [1987], Crnković and Witten [1987], and Ashtekar et al. [1991].

Examples

a Particle Mechanics. For particle mechanics Σ is a point, and τ maps Σ to some $t \in \mathbb{R}$. We identify $\mathcal{Y}_\tau$ with Q and $\mathcal{Z}_\tau$ with $\mathbb{R} \times T^*Q$, with coordinates (q^A, p, p_A). The one-form θ_τ is $\theta_\tau = p_A dq^A$ and R_τ is given by $(q^A, p, p_A) \mapsto (q^A, p_A)$. Thus the τ-phase space is just T^*Q, and the process of reducing the multisymplectic formalism to the instantaneous formalism in particle mechanics is simply reduction to the autonomous case.

b Electromagnetism. In the case of electromagnetism on a fixed background spacetime, Σ is a 3-manifold and $\tau \in \mathrm{Emb}(\Sigma, X)$ is a parametrized spacelike hypersurface. The space $\mathcal{Y}_\tau$ consists of fields A_ν over Σ_τ, $T^*\mathcal{Y}_\tau$ consists of fields and their conjugate momenta $(A_\nu, \mathfrak{E}^\nu)$ on Σ_τ, while the space $\mathcal{Z}_\tau$ consists of fields and multimomenta fields $(A_\nu, p, \mathfrak{F}^{\nu\mu})$ on Σ_τ. In adapted coordinates the map R_τ is given by

$$(A_\nu, p, \mathfrak{F}^{\nu\mu}) \mapsto (A_\nu, \mathfrak{E}^\nu), \tag{5C.6}$$

where $\mathfrak{E}^\nu = \mathfrak{F}^{\nu 0}$. The canonical momentum $\mathfrak{E}^\nu$ can thus be identified with the negative of the electric field density. The symplectic structure on $T^*\mathcal{Y}_\tau$ takes the form

$$\omega_\tau(A, \mathfrak{E}) = \int_{\Sigma_\tau} (dA_\nu \wedge d\mathfrak{E}^\nu) \otimes d^3x_0. \tag{5C.7}$$

When electromagnetism is parametrized, we simply append the metric $g_{\sigma\rho}$ and its corresponding multimomenta $\rho^{\sigma\rho\mu}$ to the other field variables as parameters. Let $S_2^{3,1}(X, \Sigma_\tau)$ denote the subspace of $S_2^{3,1}(X)_\tau$ consisting of Lorentz metrics on X relative to which Σ_τ is spacelike. Thus we replace $\mathcal{Y}_\tau$ by

$$\tilde{\mathcal{Y}}_\tau = \mathcal{Y}_\tau \times \mathcal{S}_2^{3,1}(X, \Sigma_\tau),$$

which consists of sections $(A; g)$ of $\tilde{Y}_\tau = Y_\tau \times_{\Sigma_\tau} S_2^{3,1}(X, \Sigma_\tau)$ over Σ_τ. Similarly, $T^*\tilde{\mathcal{Y}}_\tau$ and $\tilde{\mathcal{Z}}_\tau$ consist of sections $(A_\nu, \mathfrak{E}^\nu; g_{\sigma\rho}, \pi^{\sigma\rho})$ and $(A_\nu, p, \mathfrak{F}^{\nu\mu}; g_{\sigma\rho}, \rho^{\sigma\rho\mu})$ over

Σ_τ, respectively. The coordinate expression for the reduction map $\tilde{R}_\tau$ now becomes

$$(A_\nu, p, \mathfrak{F}^{\nu\mu}; g_{\sigma\rho}, \rho^{\sigma\rho\mu}) \mapsto (A_\nu, \mathfrak{E}^\nu; g_{\sigma\rho}, \pi^{\sigma\rho}) \tag{5C.8}$$

where $\pi^{\sigma\rho} = \rho^{\sigma\rho 0}$. Finally, the symplectic structure on $T^*\tilde{\mathcal{Y}}_\tau$ is

$$\tilde{\omega}_\tau(A, \mathfrak{E}; g, \pi) = \int_{\Sigma_\tau} (dA_\nu \wedge d\mathfrak{E}^\nu + dg_{\sigma\rho} \wedge d\pi^{\sigma\rho}) \otimes d^3x_0. \tag{5C.9}$$

c A Topological Field Theory. Since in a topological field theory there is no metric on X, it does not make sense to speak of "spacelike hypersurfaces" (although we shall continue to informally refer to Σ_τ as a "Cauchy surface"). Thus we may take τ to be *any* embedding of Σ into X.

Other than this, along with the fact that Σ is 2-dimensional, Chern–Simons theory is much the same as electromagnetism. Specifically, $\mathcal{Y}_\tau$ consists of fields A_ν over Σ_τ, $T^*\mathcal{Y}_\tau$ consists of fields and their conjugate momenta (A_ν, π^ν) over Σ_τ, and $\mathcal{Z}_\tau$ consists of fields and their multimomenta $(A_\nu, p, p^{\nu\mu})$ over Σ_τ. Then R_τ and ω_τ are given by

$$(A_\nu, p, p^{\nu\mu}) \mapsto (A_\nu, \pi^\nu) \tag{5C.10}$$

and

$$\omega_\tau(A, \pi) = \int_{\Sigma_\tau} (dA_\nu \wedge d\pi^\nu) \otimes d^2x_0 \tag{5C.11}$$

respectively, where $\pi^\nu = p^{\nu 0}$.

d Bosonic Strings. Here Σ is a 1-manifold and $\tau \in \mathrm{Emb}(\Sigma, X)$ is a parametrized curve in X. Now $\mathcal{Y}_\tau$ consists of sections $(\varphi^A, h_{\sigma\rho})$ of

$$(X \times M) \times_{\Sigma_\tau} S_2^{1,1}(X, \Sigma_\tau),$$

$T^*\mathcal{Y}_\tau$ consists of fields and their conjugate momenta $(\varphi^A, h_{\sigma\rho}, \pi_A, \varpi^{\sigma\rho})$, and $\mathcal{Z}_\tau$ consists of fields and their multimomenta $(\varphi^A, h_{\sigma\rho}, p, p_A{}^\mu, \rho^{\sigma\rho\mu})$, all over Σ_τ. In adapted coordinates, the map R_τ is

$$(\varphi^A, h_{\sigma\rho}, p, p_A{}^\mu, \rho^{\sigma\rho\mu}) \mapsto (\varphi^A, h_{\sigma\rho}, \pi_A, \varpi^{\sigma\rho}) \tag{5C.12}$$

where $\pi_A = p_A{}^0$ and $\varpi^{\sigma\rho} = \rho^{\sigma\rho 0}$. The symplectic form on $T^*\mathcal{Y}_\tau$ is then

$$\omega_\tau(\varphi, h, \pi, \varpi) = \int_{\Sigma_\tau} (d\varphi^A \wedge d\pi_A + dh_{\sigma\rho} \wedge d\varpi^{\sigma\rho}) \otimes d^1x_0. \tag{5C.13}$$

---

6 Initial Value Analysis of Field Theories

In the previous chapter we showed how to space + time decompose multisymplectic structures. Here we perform a similar decomposition of dynamics using the notion of slicings. This material puts the standard initial value analysis into our context, with a few clarifications concerning how to intrinsically split off the time derivatives of fields in the passage from the covariant to the instantaneous pictures. A main result of this chapter is that the dynamics is compatible with the space + time decomposition in the sense that Hamiltonian dynamics in the instantaneous formalism corresponds directly to the covariant Lagrangian dynamics of Chapter 3; see §6D. We also discuss a symplectic version of the Dirac–Bergmann treatment of degenerate Hamiltonian systems, initial value constraints, and gauge transformations in §6E.

6A Slicings

To discuss dynamics, that is, how fields evolve in time, we define a global notion of "time." This is accomplished by introducing "slicings" of spacetime and the relevant bundles over it.

A ***slicing*** of an $(n+1)$-dimensional spacetime X consists of an n-dimensional manifold Σ (sometimes known as a ***reference Cauchy surface***) and a diffeomorphism

$$\mathfrak{s}_X : \Sigma \times \mathbb{R} \to X.$$

For $\lambda \in \mathbb{R}$, we write $\Sigma_\lambda = \mathfrak{s}_X(\Sigma \times \{\lambda\})$ and $\tau_\lambda : \Sigma \to \Sigma_\lambda \subset X$ for the embedding defined by $\tau_\lambda(x) = \mathfrak{s}_X(x, \lambda)$. See Figure 6.1. The slicing parameter λ gives rise to a global notion of "time" on X which need not coincide with locally defined coordinate time, nor with proper time along the curves $\lambda \mapsto \mathfrak{s}_X(x, \lambda)$. The ***generator*** of $\mathfrak{s}_X$ is the vector field ζ_X on X defined by

$$\frac{\partial}{\partial \lambda} \mathfrak{s}_X(x, \lambda) = \zeta_X(\mathfrak{s}_X(x, \lambda)).$$

Alternatively, ζ_X is the push-forward by $\mathfrak{s}_X$ of the standard vector field $\partial/\partial\lambda$ on $\Sigma \times \mathbb{R}$; that is,

$$\zeta_X = T\mathfrak{s}_X \cdot \frac{\partial}{\partial \lambda}. \tag{6A.1}$$

Given a bundle $K \to X$ and a slicing $\mathfrak{s}_X$ of X, a ***compatible slicing*** of K is a bundle $K_\Sigma \to \Sigma$ and a bundle diffeomorphism $\mathfrak{s}_K : K_\Sigma \times \mathbb{R} \to K$ such that

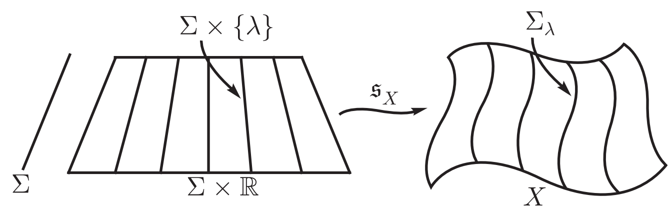

Figure 6.1: A slicing of spacetime

the diagram

$$\begin{array}{ccc} K_\Sigma \times \mathbb{R} & \xrightarrow{\mathfrak{s}_K} & K \\ \big\downarrow & & \big\downarrow \\ \Sigma \times \mathbb{R} & \xrightarrow{\mathfrak{s}_X} & X \end{array} \tag{6A.2}$$

commutes, where the vertical arrows are bundle projections. We write $K_\lambda = \mathfrak{s}_K(K_\Sigma \times \{\lambda\})$ and $\mathfrak{s}_\lambda : K_\Sigma \to K_\lambda \subset K$ for the embedding defined by $\mathfrak{s}_\lambda(k) = \mathfrak{s}_K(k, \lambda)$, as in Figure 6.2. The generating vector field ζ_K of $\mathfrak{s}_K$ is defined by a formula analogous to (6A.1). Note that ζ_K and ζ_X are complete and everywhere transverse to the slices K_λ and Σ_λ, respectively.

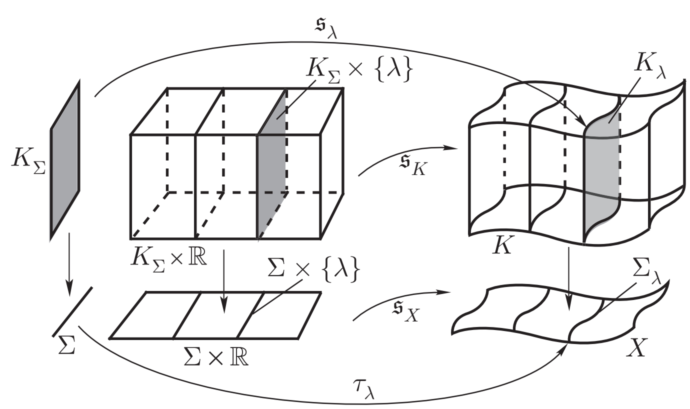

Figure 6.2: A slicing of the bundle K

Every compatible slicing $(\mathfrak{s}_K, \mathfrak{s}_X)$ of $K \to X$ defines a one-parameter group of bundle automorphisms: the flow f_λ of the generating vector field ζ_K, which is given by

$$f_\lambda(k) = \mathfrak{s}_K(\mathfrak{s}_K^{-1}(k) + \lambda),$$

where "$+\ \lambda$" means addition of λ to the second factor of $K_\Sigma \times \mathbb{R}$. This flow is fiber-preserving since ζ_K projects to ζ_X. Conversely, let f_λ be a fiber-preserving flow on K with generating vector field ζ_K. Then ζ_K along with a choice of Cauchy surface Σ_τ such that $\zeta_X \pitchfork \Sigma_\tau$ determines (at least in a neighborhood of K_τ in K) a slicing $\mathfrak{s}_K : K_\tau \times \mathbb{R} \to K$ according to $\mathfrak{s}_K(k, \lambda) = f_\lambda(k)$. Any other slicing corresponding to the above data differs from this $\mathfrak{s}_K$ by a diffeomorphism.

Slicings of bundles give rise to trivializations of associated spaces of sections. Given $K \to X$, recall from §5A that we have the bundle

$$\mathcal{K}^\Sigma = \bigcup_{\tau \in \mathrm{Emb}(\Sigma, X)} \mathcal{K}_\tau$$

over $\mathrm{Emb}(\Sigma, X)$, where $\mathcal{K}_\tau$ is the space of sections of $K_\tau = K \,\big|\, \Sigma_\tau$. Let $\mathcal{K}^\tau$ denote the portion of $\mathcal{K}^\Sigma$ that lies over the curve of embeddings $\lambda \mapsto \tau_\lambda$, where $\lambda \in \mathbb{R}$. In other words,

$$\mathcal{K}^\tau = \bigcup_{\lambda \in \mathbb{R}} \mathcal{K}_\lambda .$$

The slicing $\mathfrak{s}_K : K_\Sigma \times \mathbb{R} \to K$ induces a trivialization $\mathfrak{s}_{\mathcal{K}} : \mathcal{K}_\Sigma \times \mathbb{R} \to \mathcal{K}^\tau$ defined by

$$\mathfrak{s}_{\mathcal{K}}(\sigma_\Sigma, \lambda) = \mathfrak{s}_\lambda \circ \sigma_\Sigma \circ \tau_\lambda^{-1}. \tag{6A.3}$$

Let $\zeta_{\mathcal{K}}$ be the pushforward of $\partial/\partial\lambda$ by means of this trivialization; then from (6A.3),

$$\zeta_{\mathcal{K}}(\sigma) = \zeta_K \circ \sigma. \tag{6A.4}$$

See Figure 6.3.

Remark 6A.1. A slicing $\mathfrak{s}_X$ of X gives rise to at least one compatible slicing $\mathfrak{s}_K$ of any bundle $K \to X$, since $X \approx \Sigma \times \mathbb{R}$ is then homotopic to Σ. ♦

Remark 6A.2. In many examples, Y is a tensor bundle over X, so $\mathfrak{s}_Y$ can naturally be induced by a slicing $\mathfrak{s}_X$ of X. Similarly, in Yang–Mills theory, slicings of the connection bundle are naturally induced by slicings of the theory's principal bundle. ♦

Remark 6A.3. Slicings of the configuration bundle $Y \to X$ naturally induce slicings of certain bundles over it. For example, a slicing $\mathfrak{s}_Y$ of Y induces a slicing $\mathfrak{s}_Z$ of Z by push-forward; if ζ_Y generates $\mathfrak{s}_Y$, then $\mathfrak{s}_Z$ is generated by the canonical lift ζ_Z of ζ_Y to Z. (As a consequence, $\pounds_{\zeta_Z}\Theta = 0$.) Likewise, a slicing of J^1Y is generated by the jet prolongation $\zeta_{J^1Y} = j^1\zeta_Y$ of ζ_Y to J^1Y. ♦

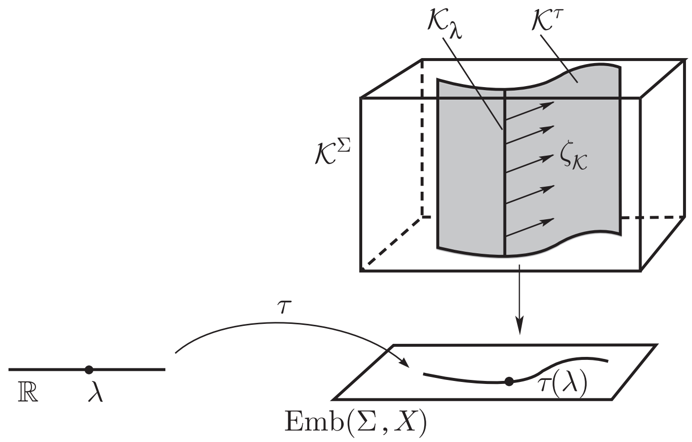

Figure 6.3: Bundles of spaces of sections

Remark 6A.4. When considering certain field theories, one may wish to modify these constructions slightly. In gravity, for example, one considers only those pairs of metrics and slicings for which each Σ_λ is spacelike. This is an open and invariant condition and so the nature of the construction is not materially changed. ♦

Remark 6A.5. It may happen that X is sufficiently complicated topologically that it cannot be globally split as $\Sigma \times \mathbb{R}$ for any Σ. In such cases one can only slice portions of spacetime and our constructions must be understood in a restricted sense. However, for globally hyperbolic spacetimes, a well-known result of Geroch (see Hawking and Ellis [1973]) states that X is homeomorphic to $\Sigma \times \mathbb{R}$, and a recent result of Bernal and Sánchez [2005] shows that in fact X is *diffeomorphic* to $\Sigma \times \mathbb{R}$. ♦

Remark 6A.6. Sometimes one wishes to allow curves of embeddings that are not slicings. (For instance, one could allow two embedded hypersurfaces to intersect.) It is known by direct calculation that the adjoint formalism (see Chapter 13) is valid even for curves of embeddings that are associated with maps $\mathfrak{s}$ that need not be diffeomorphisms. See, for example, Fischer and Marsden [1979a]. ♦

Remark 6A.7. In the instantaneous formalism, dynamics is usually studied relative to a fixed slicing of spacetime and the bundles over it. It is important to know to what extent the dynamics is the "same" for all possible slicings. To this end we introduce in Part IV fiducial models of all relevant objects which are *universal* for all slicings in the sense that one can work abstractly on the fixed model objects and then transfer the results to the spacetime context by means of a slicing. This provides a natural mechanism for comparing the results obtained by using different slicings. ♦

Remark 6A.8. In practice, the one-parameter group of automorphisms of the configuration bundle Y associated to a slicing is often induced by a one-parameter subgroup of the gauge group $\mathcal{G}$ of the theory; let us call such slicings ***$\mathcal{G}$-slicings***. In fact, later we will focus on slicings which arise in this way via the gauge group action. For $\mathcal{G}$-slicings we have $\zeta_Y = \xi_Y$ for some $\xi \in \mathfrak{g}$. This provides a crucial link between dynamics and the gauge group, and will ultimately enable us in §7D to correlate the Hamiltonian with the energy-momentum map for the gauge group action. For classical fields propagating on a fixed background spacetime, it is necessary to treat the background metric parametrically—so that $\mathcal{G}$ projects *onto* $\mathrm{Diff}(X)$—to obtain such slicings. (See Remark 8B.1.) ♦

Remark 6A.9. For some topological field theories, there is a subtle interplay between the existence of a slicing of spacetime and that of a symplectic structure on the space of solutions of the field equations. See Horowitz [1989] for a discussion. ♦

Remark 6A.10. Often slicings of X are arranged to implement certain "gauge conditions" on the fields. For example, in Maxwell's theory one may choose a slicing relative to which the Coulomb gauge condition $\nabla \cdot \mathbf{A} = 0$ holds. In general relativity, one often chooses a slicing of a given spacetime so that each hypersurface Σ_λ has constant mean curvature. This can be accomplished by solving the adjoint equations (1.3) together with the gauge conditions, which will simultaneously generate a slicing of spacetime and a solution of the field equations, with the solution "hooked" to the slicing via the gauge condition. Note that in this case the slicing is not predetermined (by specifying the atlas fields $\alpha_i(\lambda)$ in advance), but rather is determined implicitly (by fixing the $\alpha_i(\lambda)$ by means of the adjoint equations together with the gauge conditions.) ♦

Remark 6A.11. In principle slicings can be chosen arbitrarily, not necessarily according to a given *a priori* rule. For example, in numerical relativity, to achieve certain accuracy goals, one may wish to choose slicings that focus on those regions in which the fields that have been computed up to that point have large gradients, thereby effectively using the slicing to produce an adaptive numerical method. In this case, the slicing is determined "on the fly" as opposed to being fixed *ab initio*. Of course, after a piece of spacetime is constructed, the slicing produced is consistent with our definitions. ♦

For a given field theory, we say that a slicing $\mathfrak{s}_Y$ of the configuration bundle Y is ***Lagrangian*** if the Lagrangian density $\mathcal{L}$ is equivariant with respect to the one-parameter groups of automorphisms associated to the induced slicings of J^1Y and $\Lambda^{n+1}X$. Let f_λ be the flow of ζ_Y so that $j^1 f_\lambda$ is the flow of ζ_{J^1Y}; then equivariance means

$$\mathcal{L}\left(j^1 f_\lambda(\gamma)\right) = (h_\lambda^{-1})^* \mathcal{L}(\gamma) \tag{6A.5}$$

for each $\lambda \in \mathbb{R}$ and $\gamma \in J^1Y$, where h_λ is the flow of ζ_X. Throughout the rest of this book we will assume that "slicing" means "Lagrangian slicing". In practice there are usually many such slicings. For example, in tensor theories, slicings of X induce slicings of Y by pull-back; these are automatically Lagrangian as long as a metric g on spacetime is included as a field variable (either variationally or parametrically). For theories on a fixed background spacetime, on the other hand, a slicing of Y typically will be Lagrangian only if the flow generated by ζ_X consists of isometries of (X, g). Since (X, g) need not have any continuous isometries, it may be necessary to treat g parametrically to obtain Lagrangian slicings. Note that by virtue of the covariance assumption **A1**, $\mathcal{G}$-slicings are automatically Lagrangian. (See, however, Example **c** following.) This requirement will play a key role in establishing the correspondence between dynamics in the covariant and $(n+1)$-formalisms.

For certain constructions we require only the notion of an ***infinitesimal slicing*** of a spacetime X. This consists of a Cauchy surface Σ_τ along with a spacetime vector field ζ_X defined over Σ_τ which is everywhere transverse to Σ_τ. We think of ζ_X as defining a "time direction" along Σ_τ. In the same vein, an ***infinitesimal slicing*** of a bundle $K \to X$ consists of K_τ along with a vector field ζ_K on K defined over K_τ which is everywhere transverse to K_τ. The infinitesimal slicings (Σ_τ, ζ_X) and (K_τ, ζ_K) are called ***compatible*** if ζ_K projects to ζ_X; we shall always assume this is the case. See Figure 6.4.

An important special case arises when the spacetime X is endowed with a Lorentzian metric g. Fix a spacelike hypersurface $\Sigma_\tau \subset X$ and let $e_\perp$ denote the future-pointing timelike unit normal vector field on Σ_τ; then $(\Sigma_\tau, e_\perp)$ is an

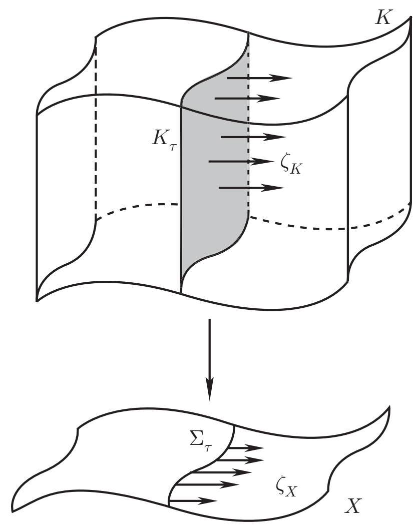

Figure 6.4: Infinitesimal slicings

infinitesimal slicing of X. In coordinates adapted to Σ_τ we expand

$$\frac{\partial}{\partial x^0} = Ne_\perp + M^i \frac{\partial}{\partial x^i}, \tag{6A.6}$$

where N is a function on Σ_τ (the ***lapse***) and $\mathbf{M} = M^i\partial/\partial x^i$ is a vector field tangent to Σ_τ (the ***shift***). It is often useful to refer an arbitrary infinitesimal slicing $\zeta_X = \zeta^\mu\partial/\partial x^\mu$ to the frame $\{e_\perp, \partial_i\}$, relative to which we have

$$\zeta_X = \zeta^0 Ne_\perp + (\zeta^0 M^i + \zeta^i)\frac{\partial}{\partial x^i}. \tag{6A.7}$$

We remark that, in general, neither $\partial/\partial x^0$ nor ζ_X need be timelike.

In both our and ADM's (Arnowitt et al. [1962]) formalisms, these lapse and shift functions play a key role. For instance, in the construction of spacetimes from initial data (say, using a computer), they are used to control the choice of slicing. This can be seen most clearly by imposing the ADM coordinate condition that $\partial/\partial x^0$ coincide with ζ_X, in which case (6A.7) reduces simply to

$$\zeta_X = Ne_\perp + \boldsymbol{M}. \tag{6A.8}$$

Examples

a Particle Mechanics. Both $X = \mathbb{R}$ and $Y = \mathbb{R} \times Q$ for particle mechanics are "already sliced" with $\zeta_X = d/dt$ and $\zeta_Y = \partial/\partial t$ respectively. From the infinitesimal equivariance equation (4D.2), it follows that this slicing is Lagrangian relative to $\mathcal{L} = L(t, q^A, v^A)dt$ iff $\partial L/\partial t = 0$, that is, L is time-independent.

One can consider more general slicings of X, interpreted as diffeomorphisms $\mathfrak{s}_X : \mathbb{R} \to \mathbb{R}$. The induced slicing $\mathfrak{s}_Y : Q \times \mathbb{R} \to Y$ given by $\mathfrak{s}_Y(q^1, \ldots, q^N, t) = (q^1, \ldots, q^N, \mathfrak{s}_X(t))$ will be Lagrangian if $\mathcal{L}$ is time reparametrization-covariant.

We can be substantially more explicit for the relativistic free particle. Consider an arbitrary slicing $Q \times \mathbb{R} \to Y$ with generating vector field

$$\zeta_Y = \chi \frac{\partial}{\partial t} + \zeta^A \frac{\partial}{\partial q^A}. \tag{6A.9}$$

From (4D.2) we see that the slicing is Lagrangian relative to (3B.8) iff

$$g_{BC,A} v^B v^C \zeta^A + g_{AC} v^C \left(\frac{\partial \zeta^A}{\partial t} + v^B \frac{\partial \zeta^A}{\partial q^B} \right) = 0. \tag{6A.10}$$

(The terms involving χ drop out as $\mathcal{L}$ is time reparametrization-covariant.) But (6A.10) holds for all $\mathbf{v}$ iff $\partial \zeta^A/\partial t = 0$ and

$$0 = g_{BC,A} v^B v^C \zeta^A + g_{AC} v^C v^B \frac{\partial \zeta^A}{\partial q^B} = v^A v^B \zeta_{(A;B)}.$$

Thus $\zeta^A \partial/\partial q^A$ must be a Killing vector field. It follows that the most general Lagrangian slicing consists of time reparametrizations horizontally and isometries vertically.

b Electromagnetism. Any slicing of the spacetime X naturally induces a slicing of the bundle $\tilde{Y} = \Lambda^1 X \times S_2^{3,1}(X)$ by push-forward. If $\zeta_X = \zeta^\mu \partial/\partial x^\mu$, the generating vector field of this induced slicing is

$$\zeta_{\tilde{Y}} = \zeta^\mu \frac{\partial}{\partial x^\mu} - A_\nu \zeta^\nu{}_{,\alpha} \frac{\partial}{\partial A_\alpha} - (g_{\sigma\mu} \zeta^\mu{}_{,\rho} + g_{\rho\mu} \zeta^\mu{}_{,\sigma}) \frac{\partial}{\partial g_{\sigma\rho}}.$$

The restriction to $\mathcal{G}$-slicings, with $\mathcal{G} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$ as in Example **b** of §4C, is not very severe for the parametrized version of Maxwell's theory.

Any complete vector field $\zeta_X = \zeta^\mu \partial/\partial x^\mu$ may be used as the generator of the spacetime slicing; then for the slicing of $\tilde{Y}$ we have the generator

$$\zeta_{\tilde{Y}} = \zeta^\mu \frac{\partial}{\partial x^\mu} + (\chi_{,\alpha} - A_\nu \zeta^\nu{}_{,\alpha}) \frac{\partial}{\partial A_\alpha} - (g_{\sigma\mu}\zeta^\mu{}_{,\rho} + g_{\rho\mu}\zeta^\mu{}_{,\sigma}) \frac{\partial}{\partial g_{\sigma\rho}}, \qquad (6A.11)$$

where χ is an arbitrary function on X (generating a Maxwell gauge transformation). A more general Lagrangian slicing (which, however, is not a $\mathcal{G}$-slicing) is obtained from this upon replacing $\chi_{,\alpha}$ by the components of a closed 1-form on X.

On the other hand, if we work with electromagnetism on a fixed spacetime background, the ζ_X must be a Killing vector field of the background metric g, and ζ_Y is of the form (6A.11) with this restriction on ζ^μ (and without the term in the direction $\partial/\partial g_{\sigma\rho}$.) If the background spacetime is Minkowskian, then ζ_X must be a generator of the Poincaré group. For a generic background spacetime, there are no Killing vectors, and hence no Lagrangian slicings. (This leads one to favor the parametrized theory.)

c A Topological Field Theory. With reference to Example **b** above, we see that with $\mathcal{G} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$, a $\mathcal{G}$-slicing of $Y = \Lambda^1 X$ is generated by

$$\zeta_Y = \zeta^\mu \frac{\partial}{\partial x^\mu} + (\chi_{,\alpha} - A_\nu \zeta^\nu{}_{,\alpha}) \frac{\partial}{\partial A_\alpha}. \qquad (6A.12)$$

Note that (6A.12) does not generate a Lagrangian slicing unless $\chi = 0$, since the replacement $A \mapsto A + \mathbf{d}\chi$ does not leave the Chern–Simons Lagrangian density invariant (cf. §4D).

d Bosonic Strings. In this case the configuration bundle

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

is already sliced with $\zeta_X = \partial/\partial x^0$ and $\zeta_Y = \partial/\partial x^0$. More generally, one can consider slicings with generators of the form

$$\zeta^\mu \frac{\partial}{\partial x^\mu} + \zeta^A \frac{\partial}{\partial \phi^A} + \zeta_{\sigma p} \frac{\partial}{\partial h_{\sigma\rho}}. \qquad (6A.13)$$

Such a slicing will be Lagrangian relative to the Lagrangian density (3B.24) iff $\zeta^A \partial/\partial\phi^A$ is a Killing vector field of (M, g) (this works much the same way as Example **a**) and

$$\zeta_{\sigma\rho} = -(h_{\sigma\alpha}\zeta^\alpha{}_{,\rho} + h_{\rho\alpha}\zeta^\alpha{}_{,\sigma}) + 2\lambda h_{\sigma\rho} \qquad (6A.14)$$

for some function λ on X. The first two terms in this expression represent that "part" of the slicing which is induced by the slicing ζ_X of X by push-forward, and the last term reflects the freedom to conformally rescale h while leaving the harmonic map Lagrangian invariant. The slicing represented by (6A.13) will be a $\mathcal{G}$-slicing, with $\mathcal{G} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X, \mathbb{R}^+)$, iff $\zeta^A = 0$.

6B Space + Time Decomposition of the Jet Bundle

In Chapter 5 we have space + time decomposed the multisymplectic formalism relative to a fixed Cauchy surface $\Sigma_\tau \in X$ to obtain the associated τ-phase space $T^*\mathcal{Y}_\tau$ with its symplectic structure $\omega_\tau = -\mathbf{d}\theta_\tau$. Now we show how to perform a similar decomposition of the jet bundle J^1Y using the notion of an infinitesimal slicing. Effectively, this enables us to invariantly separate the temporal from the spatial derivatives of the fields.

Fix an infinitesimal slicing $(Y_\tau, \zeta := \zeta_Y)$ of Y and set

$$\varphi := \phi \,\big|\, \Sigma_\tau \quad \text{and} \quad \dot{\varphi} := \pounds_\zeta \phi \,\big|\, \Sigma_\tau,$$

so that in coordinates

$$\varphi^A = \phi^A \,\big|\, \Sigma_\tau \quad \text{and} \quad \dot{\varphi}^A = (\zeta^\mu \phi^A{}_{,\mu} - \zeta^A \circ \phi) \,\big|\, \Sigma_\tau. \tag{6B.1}$$

Define an affine bundle map $\beta_\zeta : (J^1Y)_\tau \to J^1(Y_\tau) \times VY_\tau$ over Y_τ by

$$\beta_\zeta(j^1\phi(x)) = (j^1\varphi(x), \dot{\varphi}(x)) \tag{6B.2}$$

for $x \in \Sigma_\tau$. In coordinates adapted to Σ_τ, (6B.2) reads

$$\beta_\zeta(x^i, y^A, v^A{}_\mu) = (x^i, y^A, v^A{}_j, \dot{y}^A). \tag{6B.3}$$

Furthermore, if the coordinates on Y are arranged so that

$$\zeta \,\big|\, Y_\tau = \frac{\partial}{\partial x^0}, \qquad \text{then} \qquad \dot{y}^A = v^A{}_0.$$

This last observation establishes:

Proposition 6B.1. *If ζ_X is transverse to Σ_τ, then β_ζ is an isomorphism.*

The bundle isomorphism β_ζ is the ***jet decomposition map*** and its inverse the ***jet reconstruction map***. Clearly, both can be extended to maps on sections; from (6B.2) we have

$$\beta_\zeta(j^1\phi \circ i_\tau) = (j^1\varphi, \dot{\varphi}) \tag{6B.4}$$

where $i_\tau : \Sigma_\tau \to X$ is the inclusion. In fact:

Corollary 6B.2. *β_ζ induces an isomorphism of $(j^1\mathcal{Y})_\tau$ with $T\mathcal{Y}_\tau$, where $(j^1\mathcal{Y})_\tau$ is the collection of restrictions of holonomic sections of $J^1Y \to X$ to Σ_τ.*[12]

Proof. Since $\dot\varphi$ is a section of VY_τ covering φ, by (5A.1) it defines an element of $T_\varphi\mathcal{Y}_\tau$. The result now follows from the previous proposition and the comment afterwards. ■

One may wish to decompose Y, as well as J^1Y, relative to a slicing. This is done so that one works with fields that are spatially covariant rather than spacetime covariant. For example, in electromagnetism, sections of $Y = \Lambda^1 X$ are one-forms $A = A_\mu dx^\mu$ over spacetime and sections of $Y_\tau = \Lambda^1 X \,\big|\, \Sigma_\tau$ are spacetime one-forms restricted to Σ_τ. One may split

$$Y_\tau = \Lambda^1\Sigma_\tau \times_{\Sigma_\tau} \Lambda^0\Sigma_\tau, \tag{6B.5}$$

so that the instantaneous configuration space consists of spatial one-forms $\mathbf{A} = A_m dx^m$ together with spatial scalars a. The map $\Lambda^1 X \,\big|\, \Sigma_\tau \to \Lambda^1\Sigma_\tau \times_{\Sigma_\tau} \Lambda^0\Sigma_\tau$ which effects this split takes the form

$$A \mapsto (\mathbf{A}, a) \tag{6B.6}$$

where $a = i_\tau^*(\zeta_X \,\lrcorner\, A)$ and $\mathbf{A} = i_\tau^* A$.

One particular case of interest is that of a metric tensor g on X. Recall that $S_2^{n,1}(X, \Sigma_\tau)$ denotes the subbundle of $S_2^{n,1}(X)_\tau$ consisting of those Lorentz metrics on X with respect to which Σ_τ is spacelike. We may space + time split

$$S_2^{n,1}(X, \Sigma_\tau) \,\big|\, \Sigma_\tau = S_2^n(\Sigma_\tau) \times_{\Sigma_\tau} T\Sigma_\tau \times_{\Sigma_\tau} \Lambda^0\Sigma_\tau \tag{6B.7}$$

as follows (cf. §21.4 of Misner et al. [1973]). Let $e_\perp$ the the forward-pointing unit timelike normal to Σ_τ, and let $N, \mathbf{M}$ be the lapse and shift functions defined via (6A.6). Set $\gamma = i_\tau^* g$, so that γ is a Riemannian metric on Σ_τ. Then the decomposition $g \mapsto (\gamma, \mathbf{M}, N)$ with respect to the infinitesimal slicing $(\Sigma_\tau, e_\perp)$ is given by

$$g = \gamma_{jk}(dx^j + M^j dt)(dx^k + M^k dt) - N^2 dt^2$$

or, in terms of matrices,

$$\begin{pmatrix} g_{00} & g_{0i} \\ & \\ g_{i0} & g_{jk} \end{pmatrix} = \begin{pmatrix} M_k M^k - N^2 & M_i \\ & \\ M_i & \gamma_{jk} \end{pmatrix}, \tag{6B.8}$$

[12] $(j^1\mathcal{Y})_\tau$ should not be confused with the collection of holonomic sections of $J^1(Y_\tau) \to \Sigma_\tau$, since the former contains information about temporal derivatives that is not included in the latter.

where $M_i = \gamma_{ij} M^j$. This decomposition has the corresponding contravariant form

$$g^{-1} = \gamma^{jk}\partial_j\partial_k - \frac{1}{N^2}(\partial_t - M^j\partial_j)(\partial_t - M^k\partial_k)$$

or, in terms of matrices,

$$\begin{pmatrix} g^{00} & g^{0i} \\ \\ g^{i0} & g^{jk} \end{pmatrix} = \begin{pmatrix} -1/N^2 & M^i/N^2 \\ \\ M^i/N^2 & \gamma^{jk} - M^j M^k/N^2 \end{pmatrix}. \tag{6B.9}$$

Furthermore, the metric volume $\sqrt{-g}$ decomposes as

$$\sqrt{-g} = N\sqrt{\gamma}. \tag{6B.10}$$

The dynamical analysis can by carried out whether or not these splits of the configuration space are done; it is largely a matter of taste. Later, in Chapters 12 and 13 when we discuss dynamic fields and atlas fields, these types of splits will play a key role.

6C The Instantaneous Legendre Transform

Using the jet reconstruction map we may space + time split the Lagrangian as follows. Define

$$\mathcal{L}_{\tau,\zeta} : J^1(Y_\tau) \times VY_\tau \to \Lambda^n\Sigma_\tau$$

by

$$\mathcal{L}_{\tau,\zeta}(j^1\varphi(x), \dot\varphi(x)) = i_\tau^* \, \mathbf{i}_{\zeta_X}\mathcal{L}(j^1\phi(x)), \tag{6C.1}$$

where $j^1\phi \circ i_\tau$ is the reconstruction of $(j^1\varphi, \dot\varphi)$. The ***instantaneous Lagrangian*** $L_{\tau,\zeta} : T\mathcal{Y}_\tau \to \mathbb{R}$ is defined by

$$L_{\tau,\zeta}(\varphi, \dot\varphi) = \int_{\Sigma_\tau} \mathcal{L}_{\tau,\zeta}(j^1\varphi, \dot\varphi) \tag{6C.2}$$

for $(\varphi, \dot\varphi) \in T\mathcal{Y}_\tau$ (cf. Corollary 6B.2). In coordinates adapted to Σ_τ this becomes, with the aid of (6C.1) and (3A.1),

$$L_{\tau,\zeta}(\varphi, \dot\varphi) = \int_{\Sigma_\tau} L(j^1\varphi, \dot\varphi)\zeta^0 d^n x_0. \tag{6C.3}$$

The instantaneous Lagrangian $L_{\tau,\zeta}$ defines an ***instantaneous Legendre transform***

$$\mathbb{F}L_{\tau,\zeta} : T\mathcal{Y}_\tau \to T^*\mathcal{Y}_\tau; \quad (\varphi, \dot\varphi) \mapsto (\varphi, \pi) \tag{6C.4}$$

in the usual way (cf. Abraham and Marsden [1978]). In adapted coordinates

$$\pi = \pi_A \, dy^A \otimes d^n x_0$$

and (6C.4) reads

$$\pi_A = \frac{\partial \mathcal{L}_{\tau,\zeta}}{\partial \dot{y}^A}. \tag{6C.5}$$

We call

$$\mathcal{P}_{\tau,\zeta} = \operatorname{im} \mathbb{F}L_{\tau,\zeta} \subset T^*\mathcal{Y}_\tau$$

the ***instantaneous*** or ***τ-primary constraint set***.

A2 Almost Regularity. *Assume that $\mathcal{P}_{\tau,\zeta}$ is a smooth, closed, submanifold of $T^*\mathcal{Y}_\tau$ and that $\mathbb{F}L_{\tau,\zeta}$ is a submersion onto its image with connected fibers.*

Remark 6C.1. Assumption **A2** is satisfied in cases of interest. ♦

Remark 6C.2. We shall see momentarily that $\mathcal{P}_{\tau,\zeta}$ is independent of ζ. ♦

Remark 6C.3. In obtaining (6C.5) we use the fact that $\mathcal{L}$ is first order. Gotay [1991b] treats the higher order case. ♦

We now investigate the relation between the covariant and instantaneous Legendre transformations. Recall that over $\mathcal{Y}_\tau$ we have the symplectic bundle map $R_\tau : (\mathcal{Z}_\tau, \Omega_\tau) \to (T^*\mathcal{Y}_\tau, \omega_\tau)$ given by

$$\langle R_\tau(\sigma), V \rangle = \int_{\Sigma_\tau} \varphi^*(\mathbf{i}_V \sigma)$$

where $\varphi = \pi_{YZ} \circ \sigma$ and $V \in T_\varphi \mathcal{Y}_\tau$.

Proposition 6C.4. *Assume ζ_X is transverse to Σ_τ. Then the following diagram commutes:*

$$\begin{array}{ccc} (j^1\mathcal{Y})_\tau & \xrightarrow{\mathbb{F}\mathcal{L}} & \mathcal{Z}_\tau \\ {\scriptstyle \beta_\zeta}\big\downarrow & & \big\downarrow{\scriptstyle R_\tau} \\ T\mathcal{Y}_\tau & \xrightarrow[\mathbb{F}L_{\tau,\zeta}]{} & T^*\mathcal{Y}_\tau \end{array} \tag{6C.6}$$

Proof. Choose adapted coordinates in which $\partial_0 \,\big|\, Y_\tau = \zeta$. Since R_τ is given by $\pi_A = {p_A}^0 \circ \sigma$, going clockwise around the diagram we obtain

$$R_\tau\big(\mathbb{F}\mathcal{L}(j^1\phi \circ i_\tau)\big) = \frac{\partial L}{\partial {v^A}_0}(\phi^B, {\phi^B}_{,\mu})\, dy^A \otimes d^n x_0.$$

This is the same as one gets going counterclockwise, taking into account (6B.3), (6C.5) and the fact that $\mathbb{F}L_{\tau,\zeta}$ is evaluated at $\dot{\varphi}^A = {\phi^A}_{,0}$. ■

We define the ***covariant primary constraint set*** to be

$$N = \mathbb{F}\mathcal{L}(J^1Y) \subset Z$$

and with a slight abuse of notation, set

$$\mathcal{N}_\tau = \mathbb{F}\mathcal{L}\left((j^1\mathcal{Y})_\tau\right) \subset \mathcal{Z}_\tau.$$

Corollary 6C.5. *If ζ_X is transverse to Σ_τ, then*

$$R_\tau(\mathcal{N}_\tau) = \mathcal{P}_{\tau,\zeta}. \tag{6C.7}$$

In particular, $\mathcal{P}_{\tau,\zeta}$ is independent of ζ, and so can be denoted simply $\mathcal{P}_\tau$.

Proof. By Corollary 6B.2, β_ζ is onto $T\mathcal{Y}_\tau$. The result now follows from the commutative diagram (6C.6). ∎

Denote by the same symbol ω_τ the pullback of the symplectic form on $T^*\mathcal{Y}_\tau$ to the submanifold $\mathcal{P}_\tau$. When there is any danger of confusion we will write $\omega_{T^*\mathcal{Y}_\tau}$ and $\omega_{\mathcal{P}_\tau}$. In general $(\mathcal{P}_\tau, \omega_\tau)$ will be merely presymplectic. However, the fact that $\mathbb{F}L_{\tau,\zeta}$ is fiber-preserving together with the almost regularity assumption **A2** imply that $\ker\omega_\tau$ *is a regular distribution* on $\mathcal{P}_\tau$ (in the sense that it defines a subbundle of $T\mathcal{P}_\tau$).

As always, the ***instantaneous Hamiltonian*** is given by

$$H_{\tau,\zeta}(\varphi,\pi) = \langle\pi,\dot\varphi\rangle - L_{\tau,\zeta}(\varphi,\dot\varphi) \tag{6C.8}$$

and is defined only on $\mathcal{P}_\tau$. The density for $H_{\tau,\zeta}$ is denoted by $\mathfrak{H}_{\tau,\zeta}$. We remark that to determine a Hamiltonian, it is essential to specify a time direction ζ on Y. This is sensible, since the system cannot evolve without knowing what "time" is. For $\zeta_Y = \xi_Y$, where $\xi \in \mathfrak{g}$, the Hamiltonian will turn out to be the negative of the energy-momentum map induced on $\mathcal{P}_\tau$ (cf. §7D). A crucial step in establishing this relationship is the following result:

Proposition 6C.6. *Let $(\varphi,\pi) \in \mathcal{P}_\tau$. Then for any holonomic lift σ of (φ,π),*

$$H_{\tau,\zeta}(\varphi,\pi) = -\int_{\Sigma_\tau} \sigma^*(\mathbf{i}_{\zeta_Z}\Theta). \tag{6C.9}$$

Here ζ_Z is the canonical lift of ζ to Z (cf. §4B). By a ***holonomic lift*** of (φ,π) we mean any element $\sigma \in R_\tau^{-1}\{(\varphi,\pi)\} \cap \mathcal{N}_\tau$. Holonomic lifts of elements of $\mathcal{P}_\tau$ always exist by virtue of Proposition 6C.4.

Proof. We will show that (6C.9) holds on the level of densities; that is,

$$\mathfrak{H}_{\tau,\zeta}(\varphi,\pi) = -\sigma^*(\mathbf{i}_{\zeta_Z}\Theta). \tag{6C.10}$$

Using adapted coordinates, (2B.11) yields

$$\sigma^*(\mathbf{i}_{\zeta_Z}\Theta) =$$
$$\left\{(p_A{}^0 \circ \sigma)\left(\zeta^A \circ \sigma - \zeta^\mu \sigma^A{}_{,\mu}\right) + \left(p \circ \sigma + (p_A{}^\mu \circ \sigma)\sigma^A{}_{,\mu}\right)\zeta^0\right\} d^n x_0$$

for any $\sigma \in \mathcal{Z}_\tau$. Now suppose that $(\varphi,\pi) \in \mathcal{P}_\tau$, and let σ be any lift of (φ,π) to $\mathcal{N}_\tau$. Thus, there is a $\phi \in \mathcal{Y}$ with $\mathbb{F}\mathcal{L} \circ j^1\phi \circ i_\tau = \sigma$. Then, using (3A.2), (6B.1), (5C.4) and (6C.1), the above becomes

$$\sigma^*(\mathbf{i}_{\zeta_Z}\Theta) = -\pi(\dot\varphi) + L(j^1\phi)\zeta^0\, d^n x_0 = -\pi(\dot\varphi) + \mathcal{L}_{\tau,\zeta}(\varphi,\dot\varphi). \qquad \blacksquare$$

Notice that (6C.9) and (6C.10) are manifestly linear in ζ_Z. This linearity foreshadows the linearity of the Hamiltonian (1.2) in the "atlas fields" to which we alluded in the Introduction.

Examples

a Particle Mechanics. First consider a nonrelativistic particle Lagrangian of the form

$$L(q,v) = \frac{1}{2}g_{AB}(q)v^A v^B + V(q).$$

Taking $\zeta = \partial/\partial t$, the Legendre transformation gives $\pi_A = g_{AB}(q)v^B$. If $g_{AB}(q)$ is invertible for all q, then $\mathbb{F}L_t$ is onto for each t and there are no primary constraints.

For the relativistic free particle, the covariant primary constraint set $N \subset Z$ is determined by the constraints

$$g^{AB}p_A p_B = -m^2 \qquad \text{and} \qquad p = 0, \tag{6C.11}$$

which follow from (3B.10).

Now fix any infinitesimal slicing

$$\left(Y_t, \zeta = \chi\frac{\partial}{\partial t} + \zeta^A\frac{\partial}{\partial q^A}\right)$$

of Y. Then we may identify $(J^1Y)_t$ with TQ according to (6B.2); that is,

$$(q^A, v^A) \mapsto (q^A, \dot{q}^A)$$

where $\dot{q}^A = \chi v^A - \zeta^A$. The instantaneous Lagrangian (6C.2) is then

$$L_{t,\zeta}(q, \dot{q}) = -m\|\dot{\mathbf{q}} + \boldsymbol{\zeta}\| \tag{6C.12}$$

(provided we take $\chi > 0$). The instantaneous Legendre transform (6C.4) gives

$$\pi_A = \frac{m g_{AB}(\dot{q}^B + \zeta^B)}{\|\dot{\mathbf{q}} + \boldsymbol{\zeta}\|}. \tag{6C.13}$$

The t-primary constraint set is then defined by the "mass constraint"

$$g^{AB}\pi_A\pi_B = -m^2. \tag{6C.14}$$

Comparing (6C.14) with (6C.11) we verify that $\mathcal{P}_t = R_t(\mathcal{N}_t)$ as predicted by (6C.7). Using (6C.14) and (6C.8) we compute

$$H_{t,\zeta}(q, \pi) = -\zeta^A \pi_A. \tag{6C.15}$$

Looking ahead to Part III (cf. also the Introduction and Remark 6E.18, it may seem curious that $H_{t,\zeta}$ does not vanish identically, since after all the relativistic free particle is a parametrized system. This is because the slicing generated by (6A.9) is *not* a $\mathcal{G}$-slicing unless $\zeta^A = 0$, in which case the Hamiltonian *does* vanish.

b Electromagnetism. First we consider the background case. Let Σ_τ be a spacelike hypersurface locally given by $x^0 = \text{constant}$, and consider the infinitesimal Lagrangian $\mathcal{G}$-slicing (Y_τ, ζ) with ζ given by

$$\zeta_Y = \zeta^\mu \frac{\partial}{\partial x^\mu} + (\chi_{,\alpha} - A_\nu {\zeta^\nu}_{,\alpha}) \frac{\partial}{\partial A_\alpha},$$

where we assume that ζ_X is a Killing vector field for g.

We construct the instantaneous Lagrangian $L_{\tau,\zeta}$. From (6B.1) we have

$$\dot{A}_\mu = \zeta^0 D_0 A_\mu + \zeta^i D_i A_\mu - (D_\mu \chi - A_\nu D_\mu \zeta^\nu), \tag{6C.16}$$

and so (3B.13) gives in particular

$$F_{0i} = \frac{1}{\zeta^0}\left(\dot{A}_i - \zeta^k D_k A_i + D_i\chi - A_\nu D_i \zeta^\nu - \zeta^0 D_i A_0\right). \tag{6C.17}$$

Substituting this into 3B.12, (6C.3) yields

$$\begin{aligned} L_{\tau,\zeta}(A,\dot A) = \int_{\Sigma_\tau} \bigg[& \frac{1}{2\zeta^0}(g^{i0}g^{j0} - g^{ij}g^{00}) \\ & \times (\dot A_i - \zeta^k A_{i,k} + \chi_{,i} - A_\nu \zeta^\nu{}_{,i} - \zeta^0 A_{0,i}) \\ & \times (\dot A_j - \zeta^m A_{j,m} + \chi_{,j} - A_\rho \zeta^\rho{}_{,j} - \zeta^0 A_{0,j}) \\ & + g^{ik}g^{0m}(\dot A_i - \zeta^k A_{i,k} + \chi_{,i} - A_\nu \zeta^\nu{}_{,i} - \zeta^0 A_{0,i})F_{km} \\ & - \frac{1}{4} g^{ik}g^{jm}F_{ij}F_{km}\zeta^0 \bigg] \sqrt{-g}\, d^3x_0. \end{aligned} \tag{6C.18}$$

The corresponding instantaneous Legendre transformation $\mathbb{F}L_{\tau,\zeta}$ is defined by

$$\begin{aligned} \mathfrak{E}^i = \bigg(& \frac{1}{\zeta^0}\big(g^{i0}g^{j0} - g^{ij}g^{00}\big)\big(\dot A_j - \zeta^m D_m A_j + D_j\chi - A_\rho D_j \zeta^\rho - \zeta^0 D_j A_0\big) \\ & + g^{ik}g^{0m}F_{km}\bigg)\sqrt{-g} \end{aligned} \tag{6C.19}$$

and

$$\mathfrak{E}^0 = 0. \tag{6C.20}$$

This last relation is the sole primary constraint in the Maxwell theory.

Thus the τ-primary constraint set is

$$\mathcal{P}_\tau = \big\{ (A,\mathfrak{E}) \in T^*\mathcal{Y}_\tau \,\big|\, \mathfrak{E}^0 = 0 \big\}. \tag{6C.21}$$

It is clear that the almost regularity assumption **A2** is satisfied in this case, and that $\mathcal{P}_\tau$ is indeed independent of the choice of ζ as required by Corollary 6C.5. Using (3B.14) and (5C.6), one can also verify that (6C.19) and (6C.20) are consistent with the covariant Legendre transformation. In particular, the primary constraint $\mathfrak{E}^0 = 0$ is a consequence of the relation $\mathfrak{E}^\nu = \mathfrak{F}^{\nu 0}$ together with the fact that $\mathfrak{F}^{\nu\mu}$ is antisymmetric on N.

Taking (6C.20) into account, (5C.7) yields the presymplectic form

$$\omega_\tau(A,\mathfrak{E}) = \int_{\Sigma_\tau} (dA_i \wedge d\mathfrak{E}^i) \otimes d^3x_0 \tag{6C.22}$$

on $\mathcal{P}_\tau$. The Hamiltonian on $\mathcal{P}_\tau$ is obtained by solving (6C.19) for $\dot A_i$ and substituting into (6C.8). After some effort, we obtain

$$H_{\tau,\zeta}(A,\mathfrak{E}) = \int_{\Sigma_\tau}\left[\frac{\zeta^0 N}{\sqrt{\gamma}}\Big(\frac{1}{2}\gamma_{ij}\mathfrak{E}^i\mathfrak{E}^j + \frac{1}{4N^2}\gamma^{ik}\gamma^{jm}\mathfrak{F}_{ij}\mathfrak{F}_{km}\Big) + \frac{1}{N\sqrt{\gamma}}(\zeta^0 M^i + \zeta^i)\mathfrak{E}^j\mathfrak{F}_{ij} + (\zeta^\mu A_\mu - \chi)_{,i}\mathfrak{E}^i\right] d^3x_0 \tag{6C.23}$$

where we have made use of the splitting (6B.8)–(6B.10) of the metric g. Note the appearance of the combination $\zeta^\mu A_\mu - \chi$ in (6C.23). Later we will recognize this as the "atlas field" for the parametrized version of Maxwell's theory. Note also the presence of the characteristic combinations $\zeta^0 N$ and $(\zeta^0 M^i + \zeta^i)$ originating from (6A.7).

For definiteness, take (X, g) to be Minkowski spacetime $(\mathbb{R}^4, \eta)$ with $\zeta_X = \partial/\partial x^0$. Thus the slicing generator ζ_Y is replaced by

$$\zeta_Y = \frac{\partial}{\partial x^0} + \chi_{,\alpha}\frac{\partial}{\partial A_\alpha} \tag{6C.24}$$

and the Hamiltonian reduces to the familiar expression

$$H_{\tau,(1,\mathbf{0})}(A,\mathfrak{E}) = \int_{\Sigma_\tau}\left[\frac{1}{2}\mathfrak{E}_i\mathfrak{E}^i + \frac{1}{4}\mathfrak{F}_{ij}\mathfrak{F}^{ij} + (A_0 - \chi)_{,i}\mathfrak{E}^i\right] d^3x_0. \tag{6C.25}$$

Now suppose that the metric g is treated parametrically. Then we no longer need to require that ζ_X be a Killing vector field, so that $\zeta_{\tilde{Y}}$ is given by (6A.11). The computations above remain unaltered, except that we obtain additional primary constraints

$$\pi^{\sigma\rho} = 0 \tag{6C.26}$$

which arise because the Lagrangian $L_{\tau,\zeta}(A,\mathfrak{E}; g, \dot{g})$ given by (6C.18) does not depend upon the metric velocities $\dot{g}$.

c A Topological Field Theory. Let Σ_τ be any compact surface in X, and fix the Lagrangian slicing

$$\zeta = \zeta^\mu\frac{\partial}{\partial x^\mu} - A_\nu\zeta^\nu{}_{,\alpha}\frac{\partial}{\partial A_\alpha} \tag{6C.27}$$

as in Example **c** of §6A. The computations are similar those in Example **b** above. In particular, (6C.16) and (6C.17) remain valid (with $\chi = 0$). Together with

(3B.19), these yield

$$L_{\tau,\zeta}(A,\dot{A}) = \int_{\Sigma_\tau} \epsilon^{0ij}\left((\dot{A}_i - \zeta^k A_{i,k} - A_\nu \zeta^\nu{}_{,i} - \zeta^0 A_{0,i})A_j + \frac{1}{2}F_{ij}A_0\zeta^0\right) d^2x_0. \quad (6C.28)$$

The instantaneous Legendre transformation is

$$\pi^i = \epsilon^{0ij}A_j \qquad \text{and} \qquad \pi^0 = 0; \quad (6C.29)$$

compare (3B.20). In contrast to electromagnetism, *all* of these relations are primary constraints. Thus the instantaneous primary constraint set is

$$\mathcal{P}_\tau = \left\{(A,\pi) \in T^*\mathcal{Y}_\tau \mid \pi^0 = 0,\ \pi^i = \epsilon^{0ij}A_j\right\}. \quad (6C.30)$$

Again we see that the regularity assumption **A2** is satisfied. From (6C.29) and (5C.11) we obtain the presymplectic form on $\mathcal{P}_\tau$,

$$\omega_\tau(A,\pi) = \int_{\Sigma_\tau} \left(\epsilon^{0ij} dA_i \wedge dA_j\right) \otimes d^2x_0. \quad (6C.31)$$

The Chern-Simons Hamiltonian is

$$H_{\tau,\zeta}(A,\pi) = \int_{\Sigma_\tau} \epsilon^{0ij}\left(\zeta^k F_{ki}A_j - \frac{1}{2}\zeta^0 F_{ij}A_0 + (\zeta^\mu A_\mu)_{,i}A_j\right) d^2x_0, \quad (6C.32)$$

which is consistent with (6C.9).

d Bosonic Strings. Consider an infinitesimal slicing (Σ_τ, ζ) as in (6A.13), with $\zeta^A = 0$. (Here we also suppose that the pull-back of h to Σ_τ is positive-definite.) Using (6B.1) and (3B.24) the instantaneous Lagrangian turns out to be

$$\begin{aligned} L_{\tau,\zeta}(\varphi,h,\dot{\varphi},\dot{h}) = -\frac{1}{2}\int_{\Sigma_\tau} \sqrt{-h}\, g_{AB}\bigg(&\frac{1}{\zeta^0}h^{00}(\dot{\varphi}^A - \zeta^1 D\varphi^A)(\dot{\varphi}^B - \zeta^1 D\varphi^B) \\ &+ 2h^{01}(\dot{\varphi}^A - \zeta^1 D\varphi^A)D\varphi^B \\ &+ \zeta^0 h^{11} D\varphi^A D\varphi^B\bigg) d^1x_0, \end{aligned} \quad (6C.33)$$

where we have set $D\varphi^A := \varphi^A{}_{,1}$. From this it follows that the instantaneous momenta are

$$\pi_A = -\sqrt{-h}\, g_{AB}\left(\frac{1}{\zeta^0}h^{00}(\dot{\varphi}^B - \zeta^1 D\varphi^B) + h^{01}D\varphi^B\right) \quad (6C.34)$$

$$\varpi^{\sigma\rho} = 0. \quad (6C.35)$$

Thus

$$\mathcal{P}_\tau = \left\{ (\varphi, h, \pi, \varpi) \in T^*\mathcal{Y}_\tau \;\middle|\; \varpi^{\sigma\rho} = 0 \right\}. \tag{6C.36}$$

This is consistent with (3B.25) and (3B.26) via (5C.12).

A short computation then gives

$$H_{\tau,\zeta}(\varphi, h, \pi, \varpi) = \\ -\int_{\Sigma_\tau} \left(\frac{1}{2h^{00}\sqrt{-h}}\,\zeta^0(\pi^2 + D\varphi^2) + \left(\frac{h^{01}}{h^{00}}\zeta^0 - \zeta^1 \right)(\pi \cdot D\varphi) \right) d^1x_0$$

for the instantaneous Hamiltonian on $\mathcal{P}_\tau$, where we have used the abbreviations $\pi^2 := g^{AB}\pi_A\pi_B$ and $\pi \cdot D\varphi := \pi_A D\varphi^A$, etc.

If we space + time split the metric h as in (6B.8)–(6B.10), then the Hamiltonian becomes simply[13]

$$H_{\tau,\zeta}(\varphi, h, \pi, \varpi) = \\ \int_{\Sigma_\tau} \left(\frac{1}{2\sqrt{\gamma}}\zeta^0 N(\pi^2 + D\varphi^2) + (\zeta^0 M + \zeta^1)(\pi \cdot D\varphi) \right) d^1x_0. \tag{6C.37}$$

This expression should be compared with its counterparts in ADM gravity in Interlude IV and §14E, and Palatini gravity in §14B. In §12C we will identify $\zeta^0 N$ and $\zeta^0 M + \zeta^1$ as the "atlas fields" for the bosonic string.

Finally, using (6C.34) and (6C.35) in (5C.13), the presymplectic structure on $\mathcal{P}_\tau$ is

$$\omega_\tau(\varphi, h, \pi, \varpi) = \int_{\Sigma_\tau} (d\varphi^A \wedge d\pi_A) \otimes d^2x_0. \tag{6C.38}$$

---

[13] Since the Lagrangian and the momenta are metric densities of weight 1 (with respect to either h or γ), the Hamiltonian must be as well. This is evidently the case for the $\pi^2/\sqrt{\gamma}$ and $\pi \cdot D\varphi$ terms, while the $D\varphi^2/\sqrt{\gamma}$ term looks anomalous. But all actually is as it should be. To see this, go up one dimension and consider the bosonic membrane. The corresponding term in the Hamiltonian would be $\sqrt{\gamma}\,\gamma^{jk} g_{AB}\varphi^A{}_{,j}\varphi^B{}_{,k}$ which *is* a density of the appropriate weight. Dropping a dimension back down to the bosonic string, this expression reduces to

$$\sqrt{\gamma}\,\gamma^{11} g_{AB}\varphi^A{}_{,1}\varphi^B{}_{,1} = D\varphi^2/\sqrt{\gamma}.$$

6D Hamiltonian Dynamics

We have now gathered together the basic ingredients of Hamiltonian dynamics: for each Cauchy surface Σ_τ, we have the τ-primary constraint set $\mathcal{P}_\tau$, a presymplectic structure ω_τ on $\mathcal{P}_\tau$, and a Hamiltonian $H_{\tau,\zeta}$ on $\mathcal{P}_\tau$ relative to a choice of evolution direction ζ. If we think of some fixed Σ_τ as the "initial time," then fields $(\varphi, \pi) \in \mathcal{P}_\tau$ are candidate initial data for the $(n+1)$-decomposed field equations; that is, Hamilton's equations. To evolve this initial data, we slice spacetime and the bundles over it into global moments of time λ.

To this end, we regard $\mathrm{Emb}(\Sigma, X)$ as the space of all (parametrized) Cauchy surfaces in the $(n+1)$-dimensional "spacetime" X. The arena for Hamiltonian dynamics in the instantaneous or $(n+1)$-formalism is the "instantaneous primary constraint bundle" $\mathcal{P}^\Sigma$ over $\mathrm{Emb}(\Sigma, X)$ whose fiber above $\tau \in \mathrm{Emb}(\Sigma, X)$ is $\mathcal{P}_\tau$.

Fix compatible slicings $\mathfrak{s}_Y$ and $\mathfrak{s}_X$ of Y and X with generating vector fields ζ and ζ_X, respectively. As in §6A, let $\tau : \mathbb{R} \to \mathrm{Emb}(\Sigma, X)$ be the curve of embeddings defined by $\tau(\lambda)(x) = \mathfrak{s}_X(x, \lambda)$.

Let $\mathcal{P}^\tau$ denote the portion of $\mathcal{P}^\Sigma$ lying over the image of τ in $\mathrm{Emb}(\Sigma, X)$. Dynamics relative to the chosen slicing takes place in $\mathcal{P}^\tau$; we view the $(n+1)$-evolution of the fields as being given by a curve

$$c(\lambda) = (\varphi(\lambda), \pi(\lambda))$$

in $\mathcal{P}^\tau$ covering $\tau(\lambda)$. All this is illustrated in Figure 6.5.

Our immediate task is to obtain the $(n+1)$-decomposed field equations on $\mathcal{P}^\tau$, which determine the curve $c(\lambda)$. This requires setting up a certain amount of notation.

Recall from §6A that the slicing $\mathfrak{s}_Y$ of Y gives rise to a trivialization $\mathfrak{s}_{\mathcal{Y}}$ of $\mathcal{Y}^\tau$, and hence induces trivializations $\mathfrak{s}_{j^1\mathcal{Y}}$ of $(j^1\mathcal{Y})^\tau$ by jet prolongation and $\mathfrak{s}_{\mathcal{Z}}$ of $\mathcal{Z}^\tau$ and $\mathfrak{s}_{T^*\mathcal{Y}}$ of $T^*\mathcal{Y}^\tau$ by pull-back. These latter trivializations are therefore *presymplectic* and *symplectic*; that is, the associated flows restrict to presymplectic and symplectic isomorphisms on fibers respectively. Furthermore, the reduction maps $R_{\tau(\lambda)} : \mathcal{Z}_{\tau(\lambda)} \to T^*\mathcal{Y}_{\tau(\lambda)}$ intertwine the trivializations $\mathfrak{s}_{\mathcal{Z}}$ and $\mathfrak{s}_{T^*\mathcal{Y}}$ in the obvious sense.

Assume that the slicing $\mathfrak{s}_Y$ of Y is Lagrangian. From Proposition 4D.1(i) $\mathbb{F}\mathcal{L} : (j^1\mathcal{Y})^\tau \to \mathcal{Z}^\tau$, regarded as a map on sections, is equivariant with respect to the (flows corresponding to the) induced trivializations of these spaces. (Infinitesimally, this is equivalent to the statement $T\mathbb{F}\mathcal{L} \cdot \zeta_{j^1\mathcal{Y}} = \zeta_{\mathcal{Z}}$ where $\zeta_{j^1\mathcal{Y}}$

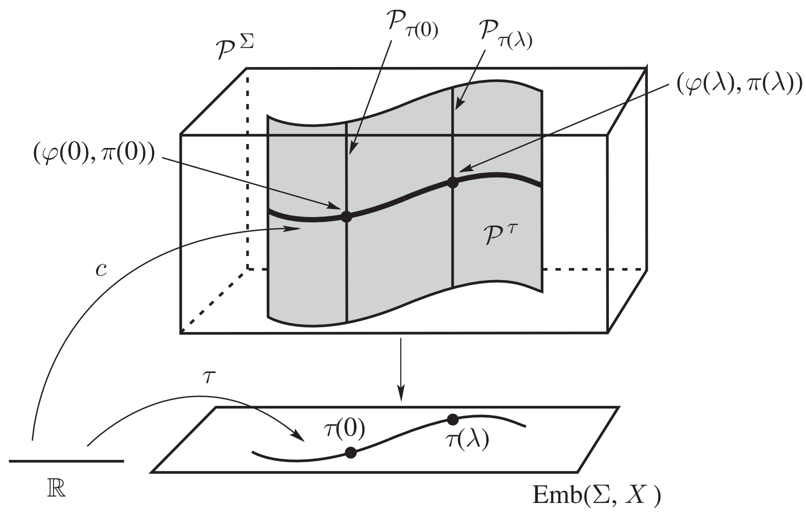

Figure 6.5: Instantaneous Dynamics

and $\zeta_{\mathcal{Z}}$ are the generating vector fields of the trivializations.) This observation, combined with the above remarks on reduction, Proposition 6C.4, and assumption **A2**, show that $\mathcal{P}^\tau$ really is a subbundle of $T^*\mathcal{Y}^\tau$, and that the symplectic trivialization $\mathfrak{s}_{T^*\mathcal{Y}}$ on $T^*\mathcal{Y}^\tau$ restricts to a presymplectic trivialization $\mathfrak{s}_{\mathcal{P}}$ of $\mathcal{P}^\tau$. We use this trivialization to coordinatize $\mathcal{P}^\tau$ by (φ, π, λ). The vector field $\zeta_{\mathcal{P}}$ which generates this trivialization is transverse to the fibers of $\mathcal{P}^\tau$ and satisfies $\zeta_{\mathcal{P}} \mathbin{\lrcorner} d\lambda = 1$. To avoid a plethora of indices (and in keeping with the notation of §6A), we will henceforth denote the fiber $\mathcal{P}_{\tau(\lambda)}$ of $\mathcal{P}^\tau$ over $\tau(\lambda) \in \mathrm{Emb}(\Sigma, X)$ simply by $\mathcal{P}_\lambda$, the presymplectic form $\omega_{\tau(\lambda)}$ by ω_λ, etc.

Using $\zeta_{\mathcal{P}}$, we may extend the forms ω_λ along the fibers $\mathcal{P}_\lambda$ to a (degenerate) 2-form ω on $\mathcal{P}^\tau$ as follows. At any point $(\varphi, \pi) \in \mathcal{P}_\lambda$, set

$$\omega(\mathcal{V}, \mathcal{W}) = \omega_\lambda(\mathcal{V}, \mathcal{W}), \tag{6D.1}$$

$$\omega(\zeta_{\mathcal{P}}, \cdot) = 0, \tag{6D.2}$$

where $\mathcal{V}, \mathcal{W}$ are vertical vectors on $\mathcal{P}^\tau$ (i.e., tangent to $\mathcal{P}_\lambda$) at (φ, π). Since $\mathcal{P}_\lambda$ has codimension one in $\mathcal{P}^\tau$, (6D.1) and (6D.2) uniquely determine ω. It is closed since ω_λ is and since the trivialization generated by $\zeta_{\mathcal{P}}$ is presymplectic (in other words, $\pounds_{\zeta_{\mathcal{P}}}\omega = 0$; cf. Gotay et al. [1983]).

Similarly, we define the function H_ζ on $\mathcal{P}^\tau$ by

$$H_\zeta(\varphi, \pi, \lambda) = H_{\lambda,\zeta}(\varphi, \pi). \tag{6D.3}$$

Tracing back through the definitions (6C.1) and (6C.2) of the instantaneous Lagrangian $L_{\lambda,\zeta}$, we find the condition that the slicing be Lagrangian guarantees that the function $L_\zeta : T\mathcal{Y}^\tau \to \mathbb{R}$ defined by

$$L_\zeta(\varphi, \dot\varphi, \lambda) = L_{\lambda,\zeta}(\varphi, \dot\varphi)$$

is independent of λ. In this case (6C.8) implies that $\zeta_{\mathcal{P}}[H_\zeta] = 0$.

Consider the 2-form $\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda$ on $\mathcal{P}^\tau$. By construction,

$$\pounds_{\zeta_{\mathcal{P}}}(\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda) = 0. \tag{6D.4}$$

We say that a curve $c : \mathbb{R} \to \mathcal{P}^\tau$ is a ***dynamical trajectory*** provided $c(\lambda)$ covers $\tau(\lambda)$ and its λ-derivative $\dot c$ satisfies

$$\dot c \,\lrcorner\, (\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda) = 0. \tag{6D.5}$$

The terminology is justified by the following result, which shows that (6D.5) is equivalent to Hamilton's equations. First note that the tangent $\dot c$ to any curve c in $\mathcal{P}^\tau$ covering τ can be uniquely split as

$$\dot c = X + \zeta_{\mathcal{P}} \tag{6D.6}$$

where X is vertical in $\mathcal{P}^\tau$. Set $X_\lambda = X \,\big|\, \mathcal{P}_\lambda$.

Proposition 6D.1. *A curve c in $\mathcal{P}^\tau$ is a dynamical trajectory iff* ***Hamilton's equations***

$$X_\lambda \,\lrcorner\, \omega_\lambda = \mathbf{d}H_{\lambda,\zeta} \tag{6D.7}$$

hold at $c(\lambda)$ for every $\lambda \in \mathbb{R}$.

Proof. With $\dot c$ as in (6D.6), we compute

$$\dot c \,\lrcorner\, (\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda) = (X \,\lrcorner\, \omega - \mathbf{d}H_\zeta) + (X[H_\zeta] + \zeta_{\mathcal{P}}[H_\zeta])\, d\lambda. \tag{6D.8}$$

A one-form α on $\mathcal{P}^\tau$ is zero iff the pull-back of α to each $\mathcal{P}_\lambda$ vanishes and $\alpha(\zeta_{\mathcal{P}}) = 0$. Applying this to (6D.8) gives

$$X_\lambda \,\lrcorner\, \omega_\lambda = \mathbf{d}H_{\lambda,\zeta}$$

which is (6D.7), and

$$-\,\zeta_{\mathcal{P}}[H_\zeta] + X[H_\zeta] + \zeta_{\mathcal{P}}[H_\zeta] = X[H_\zeta] = 0. \tag{6D.9}$$

But (6D.7) implies (6D.9), because ω_λ is skew-symmetric. ∎

Suppose $\mathcal{V}$ is a vector field on $\mathcal{P}^\tau$ whose integral curves are dynamical trajectories, so that $\pounds_{\mathcal{V}}(\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda) = 0$. Let $\mathcal{F}_{\lambda_1,\lambda_2} : \mathcal{P}_{\lambda_1} \to \mathcal{P}_{\lambda_2}$ be its flow. From the considerations above, we immediately obtain

Corollary 6D.2. $\mathcal{F}_{\lambda_1,\lambda_2}$ *is symplectic, i.e.,*

$$\mathcal{F}^*_{\lambda_1,\lambda_2}\omega_{\lambda_2} = \omega_{\lambda_1}.$$

Thus, Hamiltonian evolution in our context is by canonical transformations. A Lagrangian version of this result is given in Marsden et al. [1998].

Remark 6D.3. The difference between the two formulations (6D.5) and (6D.7) of the dynamical equations is mainly one of outlook. Equation (6D.5) corresponds to the approach usually taken in time-dependent mechanics (à la Cartan), while (6D.7) is usually seen in the context of conservative mechanics (à la Hamilton), cf. Chapters 3 and 5 of Abraham and Marsden [1978]. We use both formulations here, since (6D.5) is most easily correlated with the covariant Euler–Lagrange equations (see below), but (6D.7) is more appropriate for a study of the initial value problem (see §6E). ♦

We now relate the Euler–Lagrange equations with Hamilton's equations in the form (6D.5). This will be done by relating the 2-form $\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda$ on $\mathcal{P}^\tau$ with the 2-form $\Omega_\mathcal{L}$ on J^1Y.

Given $\phi \in \mathcal{Y}$, set $\sigma = \mathbb{F}\mathcal{L}(j^1\phi)$. Using the slicing, we map σ to a curve c_ϕ in $\mathcal{P}^\tau$ by applying the reduction map R_λ to σ at each instant λ; that is,

$$c_\phi(\lambda) = R_\lambda(\sigma_\lambda) \tag{6D.10}$$

where $\sigma_\lambda = \sigma \circ i_\lambda$ and $i_\lambda : \Sigma_\lambda \to X$ is the inclusion. (That $c_\phi(\lambda) \in \mathcal{P}_\lambda$ for each λ follows from the commutativity of diagram (6C.6).) The curve c_ϕ is called the ***canonical decomposition*** of the spacetime field ϕ with respect to the given slicing.

The main result of this section is the following, which asserts the equivalence of the Euler–Lagrange equations with Hamilton's equations.

Theorem 6D.4. *Assume* **A2**.

(*i*) *Let the spacetime field ϕ be a solution of the Euler–Lagrange equations. Then its canonical decomposition c_ϕ with respect to any Lagrangian slicing satisfies Hamilton's equations.*

(*ii*) *Conversely, every solution of Hamilton's equations is the canonical decomposition (with respect to some slicing) of a solution of the Euler–Lagrange equations.*

We observe that if ϕ is defined only locally (i.e., in a neighborhood of a Cauchy surface) and c_ϕ is defined in a corresponding interval $(a, b) \in \mathbb{R}$, then the theorem remains true.

Recall from Theorem 3B.1 that ϕ is a solution of the Euler–Lagrange equations iff

$$(j^1\phi)^*(\mathbf{i}_V \Omega_{\mathcal{L}}) = 0 \tag{6D.11}$$

for all vector fields V on J^1Y. Recall also that this statement remains valid if we require V to be π_{X,J^1Y}-vertical. Let V be any such vector field defined along $j^1\phi$ and set $W = T\mathbb{F}\mathcal{L} \cdot V$. For each $\lambda \in \mathbb{R}$, define the vector $\mathcal{W}_\lambda \in T_{c(\lambda)}\mathcal{P}_\lambda$ by

$$\mathcal{W}_\lambda = TR_\lambda \cdot (W \circ \sigma_\lambda). \tag{6D.12}$$

As λ varies, this defines a vertical vector field $\mathcal{W}$ on $\mathcal{P}^\tau$ along c_ϕ.

Lemma 6D.5. *Let V be a π_{X,J^1Y}-vertical vector field on J^1Y and $\phi \in \mathcal{Y}$. With notation as above, we have*

$$\int_{c_\phi} \mathbf{i}_{\mathcal{W}}(\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda) = \int_X (j^1\phi)^*(\mathbf{i}_V \Omega_{\mathcal{L}}). \tag{6D.13}$$

Proof. The left hand side of (6D.13) is

$$\int_{\mathbb{R}} \left\{ \mathbf{i}_{\dot{c}_\phi} \mathbf{i}_{\mathcal{W}}(\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda) \right\} d\lambda,$$

while the right hand side is

$$\int_{\Sigma \times \mathbb{R}} \mathfrak{s}_X^*(j^1\phi)^*(\mathbf{i}_V \Omega_{\mathcal{L}}) = \int_{\mathbb{R}} \left\{ \int_\Sigma \mathbf{i}_{\partial/\partial\lambda} \mathfrak{s}_X^*(j^1\phi)^*(\mathbf{i}_V \Omega_{\mathcal{L}}) \right\} d\lambda.$$

Thus, to prove (6D.13), it suffices to show that

$$(\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda)(\mathcal{W}, \dot{c}_\phi) = \int_\Sigma \mathbf{i}_{\partial/\partial\lambda} \mathfrak{s}_X^*(j^1\phi)^*(\mathbf{i}_V \Omega_{\mathcal{L}}). \tag{6D.14}$$

Using (3B.2), the right hand side of (6D.14) becomes

$$\begin{aligned}
\int_\Sigma \mathbf{i}_{\partial/\partial\lambda} \mathfrak{s}_X^*(j^1\phi)^*(\mathbf{i}_V \mathbb{F}\mathcal{L}^*\Omega) &= \int_\Sigma \mathbf{i}_{\partial/\partial\lambda} \mathfrak{s}_X^* \sigma^*(\mathbf{i}_W \Omega) = \int_\Sigma \tau_\lambda^* \left[\mathbf{i}_{\zeta_X} \sigma^*(\mathbf{i}_W \Omega) \right] \\
&= \int_{\Sigma_\lambda} i_\lambda^* \left[\mathbf{i}_{\zeta_X} \sigma^*(\mathbf{i}_W \Omega) \right] = \int_{\Sigma_\lambda} i_\lambda^* \sigma^* \left(\mathbf{i}_{T\sigma \cdot \zeta_X} \mathbf{i}_W \Omega \right) \\
&= \int_{\Sigma_\lambda} \sigma_\lambda^* \left(\mathbf{i}_{T\sigma \cdot \zeta_X} \mathbf{i}_W \Omega \right).
\end{aligned}$$

By adding and subtracting the same term, rewrite this as

$$\int_{\Sigma_\lambda} \sigma_\lambda^* \left(\mathbf{i}_{T\sigma\cdot\zeta_X - \zeta_Z}\mathbf{i}_W\Omega\right) + \int_{\Sigma_\lambda} \sigma_\lambda^* \left(\mathbf{i}_{\zeta_Z}\mathbf{i}_W\Omega\right), \tag{6D.15}$$

where ζ_Z is the generating vector field of the induced slicing of Z.

We claim that the first term in (6D.15) is equal to $\omega(\mathcal{W}, \dot{c}_\phi)$. Indeed, since $T\sigma \cdot \zeta_X - \zeta_Z$ is π_{XZ}-vertical, (5B.8) and the fact that R_λ is canonical give

$$\begin{aligned}
&\int_{\Sigma_\lambda} \sigma_\lambda^* \left(\mathbf{i}_{T\sigma\cdot\zeta_X - \zeta_Z}\mathbf{i}_W\Omega\right)\\
&\qquad = \Omega_\lambda(\sigma_\lambda)(W \circ \sigma_\lambda, (T\sigma \cdot \zeta_X - \zeta_Z) \circ \sigma_\lambda)\\
&\qquad = \omega_\lambda(c_\phi(\lambda))(TR_\lambda \cdot [W \circ \sigma_\lambda], TR_\lambda \cdot [(T\sigma \cdot \zeta_X - \zeta_Z) \circ \sigma_\lambda]).
\end{aligned}$$

Think of σ as a curve $\mathbb{R} \to \mathcal{N}^\tau \subset \mathcal{Z}^\tau$ according to $\lambda \mapsto \sigma_\lambda$. The tangent to this curve at time λ is $(T\sigma \cdot \zeta_X) \circ \sigma_\lambda$ and, from (6A.4), which states that $\zeta_{\mathcal{Z}}(\sigma) = \zeta_Z \circ \sigma$, its vertical component is thus $(T\sigma \cdot \zeta_X - \zeta_Z) \circ \sigma_\lambda$. Since the curve σ is mapped onto the curve c_ϕ by R_λ, it follows that $TR_\lambda \cdot [(T\sigma \cdot \zeta_X - \zeta_Z) \circ \sigma_\lambda]$ is the vertical component X_λ of $\dot{c}_\phi(\lambda)$. Thus in view of (6D.12), (6D.6), (6D.1), and (6D.2), the above becomes

$$\omega_\lambda(c_\phi(\lambda))(\mathcal{W}_\lambda, X_\lambda) = \omega(c_\phi(\lambda))(\mathcal{W}, \dot{c}_\phi),$$

as claimed.

Finally, we show that the second term in (6D.15) is just $\mathbf{d}H_\zeta \wedge \mathbf{d}\lambda(\mathcal{W}, \dot{c}_\phi)$. We compute at $c_\phi(\lambda) = R_\lambda(\sigma_\lambda)$:

$$\begin{aligned}
\mathbf{d}H_\zeta \wedge \mathbf{d}\lambda(\mathcal{W}, \dot{c}_\phi) &= \mathcal{W}[H_\zeta] = \mathcal{W}_\lambda[H_{\lambda,\zeta}]\\
&= -\mathcal{W}_\lambda\left[\int_{\Sigma_\lambda} \sigma_\lambda^*(\mathbf{i}_{\zeta_Z}\Theta)\right] = -\int_{\Sigma_\lambda} \sigma_\lambda^*(\pounds_W \mathbf{i}_{\zeta_Z}\Theta)
\end{aligned}$$

where we have used (6D.3), (6C.9) and (6D.12). By Stokes' theorem, this equals

$$-\int_{\Sigma_\lambda} \sigma_\lambda^*(\mathbf{i}_W \mathbf{d}\mathbf{i}_{\zeta_Z}\Theta) = -\int_{\Sigma_\lambda} \sigma_\lambda^*(\mathbf{i}_W \pounds_{\zeta_Z}\Theta) - \int_{\Sigma_\lambda} \sigma_\lambda^*(\mathbf{i}_W \mathbf{i}_{\zeta_Z}\Omega)$$

and the first term here vanishes since ζ_Z is a canonical lift (cf. Remark 6A.3). ■

Proof of Theorem 6D.4. (*i*) First, suppose that ϕ is a solution of the Euler-Lagrange equations. From Theorem 3B.1 the right hand side of (6D.14) vanishes. Thus

$$(\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda)(\mathcal{W}, \dot{c}_\phi) = 0 \tag{6D.16}$$

for all $\mathcal{W}$ given by (6D.12). By **A2** and Proposition 6C.4, every vector on $\mathcal{P}^\tau$ has the form $\mathcal{W} + f\dot{c}_\phi$ for some $\mathcal{W}$ and some function f on $\mathcal{P}^\tau$. Since the form $\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda$ vanishes on $(\dot{c}_\phi, \dot{c}_\phi)$, it follows from (6D.16) that $\dot{c}_\phi$ is in the kernel of $\omega + \mathbf{d}H_\zeta \wedge \mathbf{d}\lambda$. The result now follows from Proposition 6D.1.

(ii) Let c be a curve in $\mathcal{P}^\tau$. By Corollary 6C.5 there exists a lift σ of c to $\mathcal{N}^\tau$; we think of σ as a section of π_{XN}. It follows from (6C.6) that $\sigma = \mathbb{F}\mathcal{L}(j^1\phi)$ for some $\phi \in \mathcal{Y}$. Thus every such curve c is the canonical decomposition of some spacetime section ϕ.

If c is a dynamical trajectory, then the right hand side of (6D.13) vanishes for every π_{X,J^1Y}-vertical vector field V on J^1Y . Arguing as in the proof of Theorem 3B.1 it follows that ϕ is a solution of the Euler–Lagrange equations. ■

6E Constraint Theory

We have just established an important equivalence between solutions of Hamilton's equations as trajectories in $\mathcal{P}^\tau$ on the one hand, and solutions of the Euler–Lagrange equations as spacetime sections of Y on the other. This does *not* imply, however, that there is a dynamical trajectory through every point in $\mathcal{P}^\tau$. Nor does it imply that if such a trajectory exists it will be unique. Indeed, two of the novel features of classical field dynamics, usually absent in particle dynamics, are the presence of both constraints on the choice of Cauchy data and unphysical ("gauge") ambiguities in the resulting evolution. In fact, essentially every classical field theory of serious interest—with the exception of pure Klein–Gordon type systems—is both over- and underdetermined in these senses. Later in Part III, we shall use the energy-momentum map (defined in §7D) as a tool for understanding the constraints and gauge freedom of classical field theories. In this section we give a rapid introduction to the more traditional theory of initial value constraints and gauge transformations following Dirac [1964] as symplectically reinterpreted by Gotay et al. [1978]. An excellent general reference is the book by Sundermeyer [1982] ; see also Gotay [1979], Gotay and Nester [1979], and Isenberg and Nester [1977].

We begin by abstracting the setup for dynamics in the instantaneous formalism as presented in §§6A–6D. Let $\mathcal{P}$ be a manifold (possibly infinite-dimensional) and let ω be a presymplectic form on $\mathcal{P}$. We consider differential equations of the form

$$\dot{p} = X(p) \tag{6E.1}$$

where the vector field X satisfies

$$\mathbf{i}_X\omega = \mathbf{d}H \tag{6E.2}$$

for some given function H on $\mathcal{P}$. Finding vector field solutions X of (6E.2) is an algebraic problem at each point. When ω is symplectic, (6E.2) has a unique solution X. But when ω is presymplectic, neither existence nor uniqueness of solutions X to (6E.2) is guaranteed. In fact, X exists at a point $p \in \mathcal{P}$ iff $\mathbf{d}H(p)$ is contained in the image of the map $T_p\mathcal{P} \to T_p^*\mathcal{P}$ determined by $X \mapsto \mathbf{i}_X\omega$.

Thus one cannot expect to find globally defined solutions X of (6E.2); in general, if X exists at all, it does so only along a submanifold $\mathcal{Q}$ of $\mathcal{P}$.[14] But there is another consideration which is central to the physical interpretation of these constructions: we want solutions X of (6E.2) to generate (finite) temporal evolution of the "fields" p from the given "Hamiltonian" H via (6E.1). But this can occur on $\mathcal{Q}$ only if X is *tangent* to $\mathcal{Q}$. Modulo considerations of well-posedness (see below), this ensures that X will generate a flow on $\mathcal{Q}$ or, in other words, that (6E.1) can be integrated. This additional requirement further reduces the set on which (6E.2) can be solved.

In Gotay et al. [1978]—hereafter abbreviated by GNH—a geometric characterization of the sets on which (6E.2) has tangential solutions is presented. The characterization relies on the notion of "symplectic polar." Let $\mathcal{Q}$ be a submanifold of $\mathcal{P}$. At each $p \in \mathcal{Q}$, we define the ***symplectic polar*** $T_p\mathcal{Q}^\perp$ of $T_p\mathcal{Q}$ in $T_p\mathcal{P}$ to be

$$T_p\mathcal{Q}^\perp = \{V \in T_p\mathcal{P} \mid \omega(V, W) = 0 \quad \text{for all} \quad W \in T_p\mathcal{Q}\}\,.$$

Set

$$T\mathcal{Q}^\perp = \bigcup_{p\in\mathcal{Q}} T_p\mathcal{Q}^\perp.$$

Then GNH proves the following result, which provides the necessary and sufficient conditions for the existence of tangential solutions to (6E.2).

Proposition 6E.1. *The equation*

$$(\mathbf{i}_X\omega - \mathbf{d}H)\big|\,\mathcal{Q} = 0 \tag{6E.3}$$

possesses solutions X tangent to $\mathcal{Q}$ iff the directional derivative of H along any vector in $T\mathcal{Q}^\perp$ vanishes:

$$T\mathcal{Q}^\perp[H] = 0. \tag{6E.4}$$

[14] We suppose that all such "constraint sets" $\mathcal{Q}$ are smooth; this issue is addressed in Remark 6E.4 following.

Moreover, GNH develop a symplectic version of Dirac's "constraint algorithm" which computes the *unique maximal* submanifold $\mathcal{C}$ of $\mathcal{P}$ along which (6E.2) possesses solutions tangent to $\mathcal{C}$. This ***final constraint submanifold*** is the limit $\mathcal{C} = \bigcap_l \mathcal{P}^l$ of a string of sequentially constructed ***constraint submanifolds***

$$\mathcal{P}^{l+1} = \left\{ p \in \mathcal{P}^l \mid (T_p\mathcal{P}^l)^\perp[H] = 0 \right\} \tag{6E.5}$$

which follow from applying the consistency conditions (6E.4) to (6E.2) beginning with $\mathcal{P}^1 = \mathcal{P}$. The basic facts are as follows.

Theorem 6E.2. *Suppose* $\mathcal{C} \neq \varnothing$. *Then*

(i) *Equation* (6E.2) *is consistent, that is, there are vector fields* $X \in \mathfrak{X}(\mathcal{C})$ *such that*

$$(\mathbf{i}_X\omega - \mathbf{d}H)\,\big|\,\mathcal{C} = 0. \tag{6E.6}$$

(ii) *If* $\mathcal{Q} \subset \mathcal{P}$ *is a submanifold along which* (6E.3) *holds with* X *tangent to* $\mathcal{Q}$, *then* $\mathcal{Q} \subset \mathcal{C}$.

The following useful characterization of the maximality of $\mathcal{C}$ follows from (ii) above and Proposition 6E.1.

Corollary 6E.3. $\mathcal{C}$ *is the largest submanifold of* $\mathcal{P}$ *with the property that*

$$T\mathcal{C}^\perp[H] = 0. \tag{6E.7}$$

These results can be thought of as providing *formal* integrability criteria for equation (6E.1), since they characterize the existence of the vector field X, but do not imply that it can actually be integrated to a flow. The latter problem is a difficult analytic one, since in classical field theory (6E.1) is usually a system of hyperbolic PDEs and great care is required (in the choice of function spaces, etc.) to guarantee that there exist solutions which propagate for finite times. We shall not consider this aspect of the theory in any depth and will simply assume, when necessary, that (6E.1) is well-posed in a suitable sense. See Hawking and Ellis [1973] and Hughes et al. [1977] for some discussion of this issue. Of course, in finite dimensions (6E.1) is a system of ODEs and so integrability is automatic.

Remark 6E.4. We assume here that each of the $\mathcal{P}^l$ as well as $\mathcal{C}$ are smooth submanifolds of $\mathcal{P}$. In practice, this need *not* be the case; the $\mathcal{P}^l$ for $l > 1$ and $\mathcal{C}$ typically have quadratic singularities (see item **7** in the Introduction and Interlude IV). This does not usually present problems, at least insofar as the computation of the constraint sets and the adjoint form are concerned. (However, Cendra and Etchechoury [2005] have recently developed a means of extending the constraint algorithm to singular cases.) In such circumstances our constructions and results must be understood to hold at smooth points. We observe, in this regard, that the singular sets of the $\mathcal{P}^l$ and $\mathcal{C}$ usually have nonzero codimension therein, and that constraint sets are "varieties" in the sense that they are the closures of their smooth points. For an introduction to some of the relevant "singular symplectic geometry", see Arms et al. [1990], Sjamaar and Lerman [1991], and Ortega and Ratiu [2004]. Of course, singularities remain important for questions of linearization stability and quantization, etc. ♦

Remark 6E.5. In infinite dimensions, Proposition 6E.1 and the construction (6E.5) of the $\mathcal{P}^l$ are not valid without additional technical qualifications which we will not enumerate here. See Gotay [1979] and Gotay and Nester [1980] for the details in the general case. ♦

Remark 6E.6. The above results pertain to the *existence* of solutions to (6E.2). It is crucial to realize that solutions, when they exist, generally are not *unique*: if X solves (6E.6), then so does $X + V$ for any vector field $V \in \ker \omega \cap \mathfrak{X}(\mathcal{C})$. Thus, besides being ***overdetermined*** (signaled by a *strict* inclusion $\mathcal{C} \subset \mathcal{P}$), equation (6E.2) is also in general ***underdetermined***, signaling the presence of gauge freedom in the theory. We will have more to say about this later. ♦

We discuss one more issue in this abstract setting: the notions of first and second class constraints. We begin by recalling the classification scheme for submanifolds of presymplectic manifolds $(\mathcal{P}, \omega)$. Let $\mathcal{C} \subset \mathcal{P}$; then $\mathcal{C}$ is

(i) ***isotropic*** if $T\mathcal{C} \subset T\mathcal{C}^\perp$

(ii) ***coisotropic*** or ***first class*** if $T\mathcal{C}^\perp \neq \{0\}$ and $T\mathcal{C}^\perp \subset T\mathcal{C}$

(iii) ***second class*** if $T\mathcal{C} + T\mathcal{C}^\perp = T_{\mathcal{C}}\mathcal{P}$.

(iv) ***symplectic*** if $T\mathcal{C} \cap T\mathcal{C}^\perp = \{0\}$.

These conditions are understood to hold at every point of $\mathcal{C}$. If $\mathcal{C}$ does not happen to fall into any of these categories, then $\mathcal{C}$ is said to be ***mixed***. Note as well that the classes are not disjoint: a submanifold can be simultaneously isotropic and coisotropic, in which case $T\mathcal{C} = T\mathcal{C}^\perp$ and $\mathcal{C}$ is called ***Lagrangian***. The difference between "second class" and "symplectic" is that $T\mathcal{C} \cap T\mathcal{C}^\perp$ may be nonzero in the former case; this distinction only appears when $\mathcal{P}$ is genuinely presymplectic. An example of a second class, yet not symplectic constraint set is the Dirac manifold for Palatini gravity, cf. §14B.

From the point of view of the submanifold $\mathcal{C}$, this classification reduces to a characterization of the closed 2-form $\omega_{\mathcal{C}}$ obtained by pulling ω back to $\mathcal{C}$. Indeed,

$$\ker \omega_{\mathcal{C}} = T\mathcal{C} \cap T\mathcal{C}^{\perp}. \tag{6E.8}$$

In particular, $\mathcal{C}$ is isotropic iff $\omega_{\mathcal{C}} = 0$ and symplectic iff $\ker \omega_{\mathcal{C}} = \{0\}$. Our main interest will be in the coisotropic case.

Before proceeding, we need to establish a technical fact which will be useful later. When the ambient space $\mathcal{P}$ is symplectic, it is well-known that $(T\mathcal{C}^{\perp})^{\perp} = T\mathcal{C}$. When $\mathcal{P}$ is merely presymplectic, this is no longer true in general. Instead we have:

Lemma 6E.7. *Let $\mathcal{C} \subset \mathcal{P}$. Then*

$$(T\mathcal{C}^{\perp})^{\perp} = T\mathcal{C} + T_{\mathcal{C}}\mathcal{P}^{\perp}.$$

In particular, if $\mathcal{C}$ is coisotropic then $(T\mathcal{C}^{\perp})^{\perp} = T\mathcal{C}$.

Proof. Symplectically imbed $(\mathcal{P}, \omega)$ into a symplectic manifold $(\mathcal{M}, \varpi)$. (For instance, $(T^*P, -\mathbf{d}\theta + \pi^*\omega)$ will do, where θ is the Liouville 1-form on $T^*\mathcal{P}$ and $\pi : T^*\mathcal{P} \to \mathcal{P}$ is the projection.) Let $\vdash$ denote the symplectic polar with respect to ϖ. It is straightforward to show that $T\mathcal{C}^{\perp} = T\mathcal{C}^{\vdash} \cap T_{\mathcal{C}}\mathcal{P}$. Now

$$\begin{aligned}
(T\mathcal{C}^{\perp})^{\perp} &= (T\mathcal{C}^{\vdash} \cap T_{\mathcal{C}}\mathcal{P})^{\perp} \\
&= (T\mathcal{C}^{\vdash} \cap T_{\mathcal{C}}\mathcal{P})^{\vdash} \cap T_{\mathcal{C}}\mathcal{P} \\
&= \big((T\mathcal{C}^{\vdash})^{\vdash} + T_{\mathcal{C}}\mathcal{P}^{\vdash}\big) \cap T_{\mathcal{C}}\mathcal{P} \\
&= (T\mathcal{C} + T_{\mathcal{C}}\mathcal{P}^{\vdash}) \cap T_{\mathcal{C}}\mathcal{P} \\
&= T\mathcal{C} + T_{\mathcal{C}}\mathcal{P}^{\perp}
\end{aligned}$$

where the third equality follows from Proposition 5.3.2 in Abraham and Marsden [1978]. This establishes the main result.

Always it is the case that $T_{\mathcal{C}}\mathcal{P}^{\perp} \subset T\mathcal{C}^{\perp}$, and if $\mathcal{C}$ is coisotropic we conclude that $T_{\mathcal{C}}\mathcal{P}^{\perp} \subset T\mathcal{C}$. The desired result now follows from the above. ■

A ***constraint*** is a function $f \in \mathcal{F}(\mathcal{P})$ which vanishes on (the final constraint set) $\mathcal{C}$. The classification of constraints depends on how they relate to $T\mathcal{C}^{\perp}$. A function f which satisfies

$$T\mathcal{C}^{\perp}[f] = 0 \tag{6E.9}$$

everywhere on $\mathcal{C}$ is said to be ***first class*** relative to $\mathcal{C}$; otherwise it is ***second class***. (These definitions are due to Dirac [1964].)

Proposition 6E.8. *(i) Let f be a constraint. Then the Hamiltonian vector field X_f of f, defined by $\mathbf{i}_{X_f}\omega = \mathbf{d}f$, exists along $\mathcal{C}$ iff $T\mathcal{P}^\perp[f] \big| \mathcal{C} = 0$. If it exists, then $X_f \in \mathfrak{X}(\mathcal{C})^\perp$.*

(ii) Conversely, suppose $(\mathcal{P}, \omega)$ is symplectic. Then at every point of $\mathcal{C}$, $T\mathcal{C}^\perp$ is pointwise spanned by the Hamiltonian vector fields of constraints.

(iii) Let f be a first class constraint. Then the Hamiltonian vector field X_f of f exists along $\mathcal{C}$ and $X_f \in \mathfrak{X}(\mathcal{C}) \cap \mathfrak{X}(\mathcal{C})^\perp$.

(iv) Conversely, suppose $(\mathcal{P}, \omega)$ is symplectic. Then at every point of $\mathcal{C}$, $T\mathcal{C} \cap T\mathcal{C}^\perp$ is pointwise spanned by the Hamiltonian vector fields of first class constraints.

(v) $\mathcal{C}$ is first class iff every constraint is first class.

(vi) $\mathcal{C}$ is second class iff every effective[15] *constraint is second class.*

Proof. We follow Patrick [1985] for parts (i)–(iv).

(i) We study the equation

$$\mathbf{i}_{X_f}\omega = \mathbf{d}f \tag{6E.10}$$

at $p \in \mathcal{C}$. The first assertion follows immediately from Proposition 6E.1 upon taking $\mathcal{Q} = \mathcal{P}$. Then, if X_f exists, $\omega(X_f, T_p\mathcal{C}) = T_p\mathcal{C}[f] = 0$ as f is a constraint, whence $X_f(p) \in T_p\mathcal{C}^\perp$.

(ii) Let $V \in T_p\mathcal{C}^\perp$ and set $\alpha = \mathbf{i}_V\omega$. Fix a neighborhood U of p in $\mathcal{P}$ and a Darboux chart $\psi : (U, \omega \big| U) \to (T_p\mathcal{P}, \omega_p)$ such that

(a) $\psi(p) = 0$,

(b) $T_p\psi = id_{T_p\mathcal{P}}$ and

(c) ψ flattens $U \cap \mathcal{C}$ onto $T_p\mathcal{C}$.

Set $f = \alpha \circ \psi$ so that, by (b), $\mathbf{d}f(p) = \mathbf{i}_V\omega$. Then (c) yields

$$f(U \cap \mathcal{C}) = \alpha(\psi(U \cap \mathcal{C})) \subset \alpha(T_p\mathcal{C}) = \omega_p(V, T_p\mathcal{C})$$

[15] A constraint f is ***effective*** provided $\mathbf{d}f \,|\, \mathcal{C} \not\equiv 0$. The reason for this requirement is that if f is any constraint, then f^2 is first class.

which vanishes as $V \in T_p\mathcal{C}^\perp$. Thus f is a constraint in U and the desired globally defined constraint is then gf, where g is a suitable bump function.

(iii) Applying Proposition 6E.1 to (6E.10) along $\mathcal{C}$ and taking (6E.9) into account, we see that X_f exists and is tangent to $\mathcal{C}$. The result now follows from (i).

(iv) Let $V \in T_p\mathcal{C} \cap T_p\mathcal{C}^\perp$. We proceed as in ($ii$); it remains to show that f is first class. For any $q \in U \cap \mathcal{C}$ and $W \in T_q\mathcal{C}^\perp$,

$$\mathbf{d}f(q) \cdot W = (\alpha \circ T_q\psi) \cdot W = \omega_p(V, T_q\psi \cdot W).$$

But ψ is a symplectic map, and consequently $T_q\psi \cdot W \in T_p\mathcal{C}^\perp$ in $T_p\mathcal{P}$. Therefore, $\omega_p(V, T_q\psi \cdot W) = 0$ as $V \in T_p\mathcal{C}$. Then gf is the desired globally defined first class constraint, where g is a suitable bump function.

(v) If $\mathcal{C}$ is first class, then $T\mathcal{C}^\perp \subset T\mathcal{C}$. Applying this to a constraint f, we see that f must be first class.

For the converse, suppose there exists $v \in T_p\mathcal{C}^\perp$ which is not tangent to $\mathcal{C}$. Extend v to a vector field V on some open set $\mathcal{U}$ containing p; let F_t be its flow. By the "straightening out theorem" (Abraham and Marsden [1978], Thm. 2.1.9), we may suppose that $\mathcal{U}$ has the form $\mathcal{W} \times (-1, 1)$, where $(w, t) = F_t(w, 0)$ and $\mathcal{W}$ is chosen so as to contain $\mathcal{C} \cap \mathcal{U}$. Then $f : \mathcal{U} \mapsto \mathbb{R}$ given by $f(w, t) = t$ is a constraint, and obviously $V(p)[f] = v[f] \neq 0$, whence f is not first class, a contradiction. (One may convert f into a globally defined second class constraint by "bumping" it, as in (ii) above.)

(vi) Suppose that f is nonsingular at $p \in \mathcal{C}$; in other words, $T_p\mathcal{P}[f] \neq 0$. Since $\mathcal{C}$ is second class, this means that $\big(T_p\mathcal{C} + T_p\mathcal{C}^\perp\big)[f] \neq 0$. As f is a constraint, this implies that $T_p\mathcal{C}^\perp[f] \neq 0$ so that f is itself second class.

Going the other way, we argue by contradiction as in the proof of the converse of (v). So assume there is $v \in T_p\mathcal{P}$ which does not belong to $T_p\mathcal{C} + T_p\mathcal{C}^\perp$. Extend v to a vector field V on some open set $\mathcal{U}$ containing p; again applying the straightening out theorem, we choose $\mathcal{W}$ so as to contain $\mathcal{C} \cap \mathcal{U}$ and so that $\big(T\mathcal{C} + T\mathcal{C}^\perp\big) \,|\, \mathcal{U} \subset T_\mathcal{C}\mathcal{W}$. Then $f : \mathcal{U} \mapsto \mathbb{R}$ given by $f(w, t) = t$ is an effective constraint, but since $\mathcal{W} = f^{-1}(0)$ and $T_p\mathcal{C} + T_p\mathcal{C}^\perp \subset T_p\mathcal{W}$ we have clearly $V(p)[f] = v[f] = 0$, whence f is not second class.

■

Remark 6E.9. Strictly speaking, X_f is defined only up to elements of $\ker \omega = \mathfrak{X}(\mathcal{P})^\perp$, but we abuse the language and continue to speak of "the" Hamiltonian vector field X_f of the constraint f. ♦

From the preceding proposition, it follows that a second class submanifold can be locally described by the vanishing of second class constraints. Similarly, if $\mathcal{C}$ is coisotropic, then all constraints are first class. In general, a mixed or isotropic submanifold will require both classes of constraints for its local description.

We now apply the abstract theory of constraints, as just described, to the study of classical field theories. To place these results into the context of dynamics in the instantaneous formalism, we fix an infinitesimal slicing (Y_τ, ζ). Then $(\mathcal{P}, \omega)$ is identified with the primary constraint submanifold $(\mathcal{P}_\tau, \omega_\tau)$ of §6C, H with the Hamiltonian $H_{\tau,\zeta}$ and (6E.2) with Hamilton's equations

$$\mathbf{i}_X \omega_\tau = \mathbf{d}H_{\tau,\zeta}, \tag{6E.11}$$

cf. §6D. We have the sequence of constraint submanifolds

$$\mathcal{C}_{\tau,\zeta} \subset \cdots \subset \mathcal{P}^l_{\tau,\zeta} \subset \cdots \subset \mathcal{P}_\tau \subset T^*\mathcal{Y}_\tau. \tag{6E.12}$$

A priori, for $l \geq 2$ the $\mathcal{P}^l_{\tau,\zeta}$ depend upon the evolution direction ζ through the consistency conditions (6E.5), as $H_{\tau,\zeta}$ does. We will soon see, however, that the final constraint set is independent of ζ.[16] As indicated in Remark 6E.4 above, for simplicity we always suppose that

A3 Regularity. *$\mathcal{C}_{\tau,\zeta}$ is a smooth manifold and $\ker \omega_{\mathcal{C}_{\tau,\zeta}} = T\mathcal{C}_{\tau,\zeta} \cap T\mathcal{C}^\perp_{\tau,\zeta}$ is a subbundle of $T\mathcal{P}_\tau | \mathcal{C}_{\tau,\zeta}$.*

To ensure that the first of these assumptions is satisfied in appropriate Sobolev spaces, one supposes that the constraints are elliptic. This issue is discussed further in Interlude IV.

The functions whose vanishing defines $\mathcal{P}_\tau$ in $T^*\mathcal{Y}_\tau$ are called ***primary constraints***; they arise because of the degeneracy of the Legendre transform. Similarly, the functions whose vanishing defines $\mathcal{P}^l_{\tau,\zeta}$ in $\mathcal{P}^{l-1}_{\tau,\zeta}$ are called ***l-ary constraints*** (secondary, tertiary, ...). These constraints are generated by the

[16] In fact, none of the $\mathcal{P}^l_{\tau,\zeta}$ depend upon ζ, but we shall not prove this here. We have already shown in Corollary 6C.5 that the primary constraint set is independent of ζ.

constraint algorithm. Sometimes, for brevity, we shall refer to all l-ary constraints for $l \geq 2$ as "secondary." When we refer to the "class" of a constraint, we will adhere to the following conventions, unless otherwise noted. The class of a constraint will always be computed relative to the final constraint set $\mathcal{C}_{\tau,\zeta}$. A *secondary* constraint f is then first class provided $T\mathcal{C}_{\tau,\zeta}^{\perp}[f] = 0$, and second class otherwise, where the polar "$\perp$" is taken with respect to $(\mathcal{P}_\tau, \omega_\tau)$. However, a *primary* constraint f is first class iff $T\mathcal{C}_{\tau,\zeta}^{\vdash}[f] = 0$, where now the polar "$\vdash$" is taken with respect to $T^*\mathcal{Y}_\tau$ with its canonical symplectic form. Similarly, if $\mathcal{Q}_\tau \subset \mathcal{P}_\tau$, the polar $T\mathcal{Q}_\tau^{\perp}$ will be taken with respect to $(\mathcal{P}_\tau, \omega_\tau)$; in particular, $\mathcal{Q}_\tau$ is coisotropic, etc., if it is so relative to the primary constraint submanifold.

These constraints are all ***initial value constraints***. Indeed, thinking of Σ_τ as the "initial time," elements $(\varphi, \pi) \in \mathcal{C}_{\tau,\zeta}$ represent admissible initial data for the $(n+1)$-decomposed field equations (6E.11). Pairs (φ, π) which do not lie in $\mathcal{C}_{\tau,\zeta}$ cannot be propagated, even formally, a finite time into the future. The next series of results will serve to make these observations precise.

Let Sol denote the set of all spacetime solutions of the Euler–Lagrange equations. (Without loss of generality, we will suppose in the rest of this section that such solutions are globally defined.) Fix a Lagrangian slicing with parameter λ. Referring back to §6D, we define a map $\mathrm{can} : \mathrm{Sol} \to \Gamma(\mathcal{P}^\tau)$ by assigning to each $\phi \in \mathrm{Sol}$ its canonical decomposition c_ϕ with respect to the slicing. Observe that, for each fixed $\lambda \in \mathbb{R}$, $\mathrm{can}_\lambda(\phi) = c_\phi(\lambda) \in \mathcal{P}_\lambda$ depends only upon ϕ and the Cauchy surface Σ_λ, but *not* on the slicing.

Proposition 6E.10. *Assume* **A2**. *Then, for each* $\lambda \in \mathbb{R}$*,*

$$\mathrm{can}_\lambda(\mathrm{Sol}) \subset \mathcal{C}_{\lambda,\zeta}.$$

Proof. Let $\phi \in \mathrm{Sol}$ and set $\lambda = 0$ for simplicity. We will show that $\mathrm{can}_0(\phi) = c_\phi(0) \in \mathcal{C}_{0,\zeta}$. Define a curve $\gamma : \mathbb{R} \to \mathcal{P}_0$ by

$$\gamma(s) = f_{-s}(c_\phi(s)) \tag{6E.13}$$

where f_s is the flow of $\zeta_{\mathcal{P}}$. We may think of c_ϕ in $\mathcal{P}^\tau$ as "collapsing" onto γ in $\mathcal{P}_0$ as in Figure 6.6.

Define a one-parameter family of curves $c^s : \mathbb{R} \to \mathcal{P}^\tau$ by

$$c^s(t) = f_{-s}(c_\phi(s+t)).$$

By Theorem 6D.4(i), c_ϕ is a dynamical trajectory. Using (6D.4) we see from (6D.5) that each curve c^s is also a dynamical trajectory "starting" at $c^s(0) = \gamma(s)$.

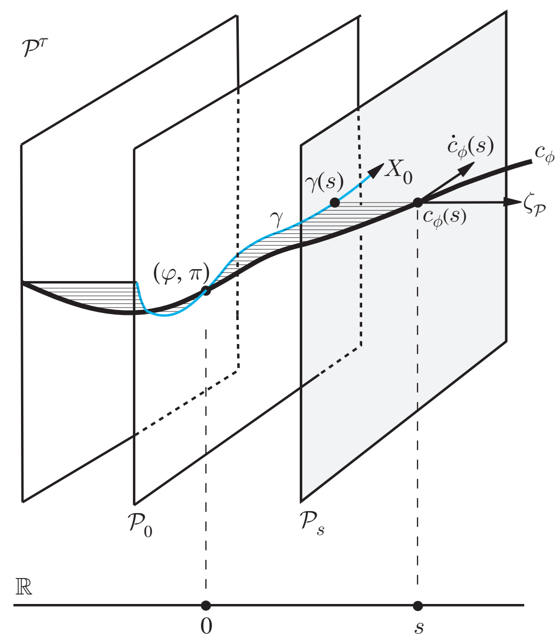

Figure 6.6: Collapsing dynamical trajectories

The tangent to each curve $c^s(t)$ at $t = 0$ takes the form

$$\frac{d}{dt} c^s(t)\bigg|_{t=0} = X_0(\gamma(s)) + \zeta_{\mathcal{P}}(\gamma(s)),$$

where X_0 is a vertical vector field on $\mathcal{P}_0$ along γ. From (6E.13) it follows that $X_0(\gamma(s))$ is the tangent to γ at s.

Proposition 6D.1 applied to each dynamical trajectory c^s at $t = 0$ implies that $X_0(\gamma(s))$ satisfies Hamilton's equations (6E.11) at each point $\gamma(s)$. Since X_0 is tangent to γ, Theorem 6E.2(ii) shows that the image of γ lies in $\mathcal{C}_{0,\zeta}$. In particular, $\gamma(0) = c_\phi(0) \in \mathcal{C}_{0,\zeta}$. ■

This proposition shows that *only* initial data $(\varphi, \pi) \in \mathcal{C}_{\lambda,\zeta}$ can be extended to solutions of the Euler–Lagrange equations. The converse is true if we assume well-posedness. We say that the Euler–Lagrange equations are ***well-posed*** relative to a slicing $\mathfrak{s}_Y$ if every $(\varphi, \pi) \in \mathcal{C}_{\lambda,\zeta}$ can be extended to a dynamical trajectory $c :]\lambda - \varepsilon, \lambda + \varepsilon[\subset \mathbb{R} \to \mathcal{P}^\tau$ with $c(\lambda) = (\varphi, \pi)$ and that this solution trajectory depends continuously (in a chosen function space topology) on (φ, π).

This will be a standing assumption in what follows.

A4 Well-Posedness *The Euler–Lagrange equations are well-posed.*

In this notion of well-posedness, one has to keep in mind that we are assuming that there is a *given* slicing of the configuration bundle Y. However, we will later prove (in Chapter 13 that well-posedness relative to one slicing with a given Cauchy surface Σ as a slice will imply well-posedness relative to any other (appropriately smooth) slicing also containing Σ as a slice.

Well-posedness for theories without gauge freedom reduces, in specific examples, to the well-posedness of a system of PDE's describing that theory in a given slicing. These will be the Hamilton equations that we have developed, written out in coordinates. In the case of metric field theories, one typically would then use theorems on strictly hyperbolic (or symmetric hyperbolic) systems to establish well-posedness (relative to a slicing by spacelike hypersurfaces); see, for example, Chernoff and Marsden [1974], John [1982], McOwen [2003], and Grundlach and Martín-García [2004].

The situation for theories with gauge freedom is a bit more subtle, because switching gauges can mix evolution and constraint equations. However, it has been established that well-posedness holds for "standard" theories such as Maxwell, Einstein, Yang-Mills and their couplings. Here, very briefly, is how the argument goes for the case of the Einstein equations (in the ADM formulation). To follow this argument, the reader will need to be familiar with works on the initial value formulation of Einstein's theory, such as Choquet-Bruhat [1962], Fischer and Marsden [1979b], Fritelli and Reula [1996], Klainerman and Nicoló [1999], Andersson and Moncrief [2002], and Choquet-Bruhat [2004].

If one has a slicing $\mathfrak{s}_Y$ specified, and one gives initial data $(\varphi, \pi) \in \mathcal{C}_{0,\zeta}$ over a Cauchy surface Σ_0, then one first takes this data and evolves it using a particular gauge or coordinate choice in which the evolution equations form a strictly hyperbolic (or symmetric hyperbolic) system.[17] This then generates a piece of spacetime on a tubular neighborhood U of the initial hypersurface and the solution ϕ so constructed (in this case the metric) on this piece of spacetime varies continuously with the choice of initial data. The solution then satisfies the Euler–Lagrange equation. Since Σ_0 is compact, there exists an $\epsilon > 0$ such that $\mathfrak{s}_X(] - \epsilon, \epsilon[\times \Sigma) \subset U$. Thus ϕ induces the required dynamical trajectory $c_\phi :]-\epsilon, \epsilon[\to \mathcal{P}^\tau$ with $c_\phi(0) = (\varphi, \pi)$ relative to the given slicing. The argument for other field theories follows a similar pattern.

[17] We caution that hyperbolicity is a gauge-dependent notion.

As was indicated in the Introduction, the above notion of well-posedness is not the same as the question of generating solutions of the initial value problem for a given choice of lapse and shift (or their generalization, called atlas fields, to other field theories) on a Cauchy surface. This is a more subtle question that we shall address later in Chapter 13. The essential difference is that with a given initial choice of lapse and shift, one still needs to construct the slicing, whereas in the present context we are assuming that a slicing has been given.

There is evidence that well-posedness fails in both of the above senses for many $R + R^2$ theories of gravity, as well as for most couplings of higher-spin fields to Einstein's theory (with supergravity being a notable exception; see Bao, Choquet-Bruhat, Isenberg, and Yasskin [1985]).

This assumption together with Proposition 6E.10 yield:

Corollary 6E.11. *If* **A2**–**A3** *hold, then* $\operatorname{can}_\lambda(\mathrm{Sol}) = \mathcal{C}_{\lambda,\zeta}$.

Since, as noted previously, $\operatorname{can}_\lambda$ depends only upon the Cauchy surface Σ_λ, we have:

Corollary 6E.12. $\mathcal{C}_{\lambda,\zeta}$ *is independent of* ζ.

Henceforth we denote the final constraint set simply by $\mathcal{C}_\lambda$. In particular, this implies that the constraint algorithm computes the same final constraint set regardless of which Hamiltonian $H_{\lambda,\zeta}$ is employed, as the generator ζ ranges over all compatible slicings (with Σ_λ as a slice).

Proposition 6E.10 shows that every dynamical trajectory $c : \mathbb{R} \to \mathcal{P}^\tau$ "collapses" to an integral curve of Hamilton's equations in $\mathcal{C}_\lambda$ for each λ. We now prove the converse; that is, every integral curve of Hamilton's equations on $\mathcal{C}_\lambda$ "suspends" to a dynamical trajectory in $\mathcal{P}^\tau$.

Proposition 6E.13. *Let γ be an integral curve of a tangential solution X_λ of Hamilton's equations on $\mathcal{C}_\lambda$. Then $c : \mathbb{R} \to \mathcal{P}^\tau$ defined by*

$$c(s) = f_s(\gamma(s)) \tag{6E.14}$$

is a dynamical trajectory.

Proof. Again setting $\lambda = 0$, (6E.14) yields

$$\dot{c}(s) = X_s(c(s)) + \zeta_{\mathcal{P}}(c(s)) \tag{6E.15}$$

where $X_s = Tf_s \cdot X_0$. Since $X_0(\gamma(s))$ satisfies (6D.7) with $\lambda = 0$ for every s, (6D.4) implies that $X_s(c(s))$ satisfies (6D.7) for every s. The desired result now follows from (6E.15) and Proposition 6D.1. ∎

Combining the proof of Proposition 6E.10 with Proposition 6E.13, we have:

Corollary 6E.14. *The Euler–Lagrange equations are well-posed iff every tangential solution X_λ of Hamilton's equations on $\mathcal{C}_\lambda$ integrates to a* (*local in time*) *flow for every $\lambda \in \mathbb{R}$.*

It remains to discuss the role of gauge transformations in constraint theory. Just as initial value constraints reflect the overdetermined nature of the field equations, gauge transformations arise when these equations are underdetermined.

Classical field theories typically exhibit ***gauge freedom*** in the sense that a given set of initial data $(\varphi, \pi) \in \mathcal{C}_\lambda$ does not suffice to uniquely determine a dynamical trajectory. Indeed, as noted in Remark 6E.6, if X_λ is a tangential solution of Hamilton's equations

$$(X_\lambda \,\lrcorner\, \omega_\lambda - \mathbf{d}H_{\lambda,\zeta})\,\big|\,\mathcal{C}_\lambda = 0, \tag{6E.16}$$

then so is $X_\lambda + V$ for any vector field $V \in \ker \omega_\lambda \cap \mathfrak{X}(\mathcal{C}_\lambda)$. For this reason we call vectors in $\ker \omega_\lambda \cap T\mathcal{C}_\lambda$ ***kinematic*** directions.

This is not the entire story, however; the indeterminacy in the solutions to the field equations is more subtle than (6E.16) would suggest. It turns out that solutions of (6E.16) are fixed only up to vector fields in $\mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$ which, in general, is larger than $\ker \omega_\lambda \cap \mathfrak{X}(\mathcal{C}_\lambda)$:

$$\ker \omega_\lambda \cap \mathfrak{X}(\mathcal{C}_\lambda) = \mathfrak{X}(\mathcal{P}_\lambda)^\perp \cap \mathfrak{X}(\mathcal{C}_\lambda) \subset \mathfrak{X}(\mathcal{C}_\lambda)^\perp \cap \mathfrak{X}(\mathcal{C}_\lambda).$$

To see this, consider a Hamiltonian vector field $V \in \mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$; according to the proof of Proposition 6E.8(*iv*), $\mathbf{i}_V \omega_\lambda = \mathbf{d}f$ where f is a first class constraint. Setting $X'_\lambda = X_\lambda + V$, (6E.16) yields

$$(X'_\lambda \,\lrcorner\, \omega_\lambda - \mathbf{d}(H_{\lambda,\zeta} + f))\,\big|\,\mathcal{C}_\lambda = 0. \tag{6E.17}$$

Thus if X_λ is a tangential solution of Hamilton's equations along $\mathcal{C}_\lambda$ with Hamiltonian $H_{\lambda,\zeta}$, then X'_λ is a tangential solution of Hamilton's equations along $\mathcal{C}_\lambda$ with Hamiltonian $H'_{\lambda,\zeta} = H_{\lambda,\zeta} + f$.

Physically, equations (6E.16) and (6E.17) are indistinguishable. Put another way, dynamics is insensitive to a modification of the Hamiltonian by the addition of a first class constraint. The reason is that $H'_{\lambda,\zeta} = H_{\lambda,\zeta}$ along $\mathcal{C}_\lambda$ and it is only what happens along $\mathcal{C}_\lambda$ that matters for the physics; distinctions that are only manifested "off" $\mathcal{C}_\lambda$—that is, in a dynamically inaccessible region—have no significance whatsoever. Thus the ambiguity in the solutions of Hamilton's equations is parametrized by $\mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$. For further discussion of these points see GNH, Gotay and Nester [1979], and Gotay [1979, 1983].

Remark 6E.15. We may rephrase the content of the last paragraph by saying that what is really of central importance for dynamics is not Hamilton's equations per se, but rather their *pullback* to $\mathcal{C}_\lambda$; the pullbacks of (6E.16) and (6E.17) to $\mathcal{C}_\lambda$ coincide. Furthermore, $\mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$ is just the kernel of the pullback of ω_λ to $\mathcal{C}_\lambda$, cf. (6E.8). ♦

Remark 6E.16. Notice also that since f is first class, (6E.7) and (6E.9) guarantee that the constraint algorithm computes the same final constraint submanifold using either Hamiltonian $H_{\lambda,\zeta}$ or $H'_{\lambda,\zeta}$. ♦

Remark 6E.17. The addition of first class constraints to the Hamiltonian (with Lagrange multipliers) is a familiar feature of the Dirac–Bergmann constraint theory. ♦

The (regular) distribution $\mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$ on $\mathcal{C}_\lambda$ is involutive and so defines a foliation of $\mathcal{C}_\lambda$. Initial data (φ, π) and (φ', π') lying on the same leaf of this foliation are said to be ***gauge-equivalent***; solutions obtained by integrating gauge-equivalent initial data cannot be distinguished physically. We call $\mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$ the ***gauge algebra*** and elements thereof ***gauge vector fields***. The flows of such vector fields preserve this foliation and hence map initial data to gauge-equivalent initial data; they are therefore referred to as ***gauge transformations***.

Proposition 6E.8 establishes the fundamental relation between gauge transformations and initial value constraints: *first class constraints generate gauge transformations*. This encapsulates a curious feature of classical field theory: the field equations being simultaneously overdetermined and underdetermined. These phenomena—*a priori* quite different and distinct—are intimately correlated via the symplectic structure. Only in special cases (i.e., when $\mathcal{C}_\lambda$ is symplectic) can the field equations be overdetermined without being underdetermined (see §8A). Conversely, it is not possible to have gauge freedom without initial value constraints.

Remark 6E.18. In Part III we will prove that the Hamiltonian (relative to a $\mathcal{G}$-slicing) of a parametrized field theory in which all fields are variational vanishes on the final constraint set. Pulling (6E.16) back to $\mathcal{C}_\lambda$ (cf. Remark

6E.15), it follows that $X_\lambda \in \ker \omega_{\mathcal{C}_\lambda}$—that is, *the evolution is totally gauge!* We will explicitly verify this in Examples **a, c** and **d** forthwith. ♦

A more detailed analysis using Proposition 6E.8 (see also Chapters 10 and 12) shows that the first class *primary* constraints correspond to gauge vector fields in $\ker \omega_\lambda \cap \mathfrak{X}(\mathcal{C}_\lambda)$, while first class *secondary* constraints correspond to the remaining gauge vector fields in $\mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$, cf. GNH. In this context, it is worthwhile to mention that second class constraints bear no relation to gauge transformations at all. For if f is second class, then by Proposition 6E.8, if it exists its Hamiltonian vector field $X_f \in T\mathcal{C}_\lambda^\perp$ everywhere along $\mathcal{C}$, but $X_f \notin T\mathcal{C}_\lambda$ at least at one point. Thus X_f tends to flow initial data *off* $\mathcal{C}_\lambda$, and hence does not generate a transformation of $\mathcal{C}_\lambda$. An extensive discussion of second class constraints is given by Lusanna [1991].

The field variables conjugate to the first class primary constraints have a special property which will be important later. We sketch the basic facts here and refer the reader to Part IV for further discussion.

Consider a nonsingular first class primary constraint f. Let g be canonically conjugate to f in the sense that

$$\omega_{T^*\mathcal{Y}_\lambda}(X_f, X_g) = 1.$$

As in the proof of the Darboux theorem, after a canonical change of coordinates, if necessary, we may write

$$\omega_{T^*\mathcal{Y}_\lambda} = \int_{\Sigma_\lambda} [\mathbf{d}g \wedge \mathbf{d}f + \cdots] \otimes d^n x_0. \tag{6E.18}$$

Expressing the evolution vector field X_λ in the form

$$X_\lambda = \frac{dg}{d\lambda}\frac{\delta}{\delta g} + \cdots$$

and substituting into Hamilton's equations (6E.16), we see that since the first term in (6E.18) vanishes when pulled back to $\mathcal{P}_\lambda$ Hamilton's equations place no restriction on $dg/d\lambda$. Thus, the evolution of g is completely arbitrary; i.e., g is purely "kinematic." Notice also from (6E.18) that

$$\frac{\delta}{\delta g} = X_f \in \ker \omega_\lambda \cap \mathfrak{X}(\mathcal{C}_\lambda),$$

which shows that $\delta/\delta g$ is a kinematic direction as defined above.

This concludes our introduction to constraint theory. In Part III we will see how both the initial value constraints and the gauge transformations can be obtained "all at once" from the energy-momentum map for the gauge group.

Examples

In the above discussion, we have treated secondary constraints (for example) as functions $f : \mathcal{P}_\lambda \to \mathbb{R}$. But in field theory, it is often more convenient (and economical) to think of them as maps $\Phi : \mathcal{P}_\lambda \to \mathcal{F}(\Sigma_\lambda)$ according to:

$$f_h(\sigma) = \int_{\Sigma_\lambda} h(\Phi \circ \sigma)\, d^n x_0$$

for each test function $h \in \mathcal{F}(\Sigma_\lambda)$. Thus the vanishing of each such Φ is equivalent to the vanishing of the "$1 \times \infty^n$" constraints f_h. We will often blur the distinction between these two interpretations.

a Particle Mechanics. We work out the details of the constraint algorithm for the relativistic free particle. Now $\mathcal{P}_\lambda \subset T^*\mathcal{Y}_\lambda$ is defined by the mass constraint (6C.14):

$$\mathfrak{H} = g^{AB}\pi_A\pi_B + m^2 = 0.$$

Then $\mathfrak{X}(\mathcal{P}_\lambda)^\perp = \ker \omega_\lambda$ is spanned by the $\omega_{T^*\mathcal{Y}_\lambda}$-Hamiltonian vector field

$$X_{\mathfrak{H}} = 2g^{AB}\pi_A \frac{\partial}{\partial q^B} - g^{AB}{}_{,C}\pi_A\pi_B \frac{\partial}{\partial \pi_C} \tag{6E.19}$$

of the "superhamiltonian" $\mathfrak{H}$. For the Hamiltonian (6C.15), the consistency conditions (6E.4) (cf. (6E.5) with $l = 1$) reduce to requiring that $X_{\mathfrak{H}}[H_{\lambda,\zeta}] = 0$. A computation gives

$$X_{\mathfrak{H}}[H_{\lambda,\zeta}] = (g^{AB}{}_{,C}\zeta^C - 2g^{AC}\zeta^B{}_{,C})\pi_A\pi_B = -2\zeta^{(A;B)}\pi_A\pi_B$$

which vanishes by virtue of the fact that the slicing is Lagrangian, so that $\zeta^A \partial/\partial q^A$ is a Killing vector field, cf. Example **a** of §6A. Thus there are no secondary constraints and so $\mathcal{C}_\lambda = \mathcal{P}_\lambda$. The mass constraint is first class.

The most general evolution vector field satisfying Hamilton's equations (6E.16) along $\mathcal{P}_\lambda$ is $X_\lambda = X + kX_{\mathfrak{H}}$, where X is any particular solution and $k \in \mathcal{F}(\mathcal{C}_\lambda)$ is arbitrary. Explicitly, writing

$$X_\lambda = \left(\frac{dq^A}{d\lambda}\right)\frac{\partial}{\partial q^A} + \left(\frac{d\pi_A}{d\lambda}\right)\frac{\partial}{\partial \pi_A},$$

the space + time decomposed equations of motion take the form

$$\begin{aligned} \frac{dq^A}{d\lambda} &= -\zeta^A + 2kg^{AB}\pi_B \\ \frac{d\pi_A}{d\lambda} &= \zeta^B{}_{,A}\pi_B - kg^{BC}{}_{,A}\pi_B\pi_C. \end{aligned} \tag{6E.20}$$

These equations appear complicated because we have written them relative to an arbitrary (but Lagrangian) slicing. If we were to choose the standard slicing $Y = Q \times \mathbb{R}$, then $\zeta^A = 0$ and (6E.20) are then clearly identifiable as the geodesic equations on (Q, g) with an arbitrary parametrization.

Since the equations of motion (6E.20) for the relativistic free particle are *ordinary* differential equations, this example is well-posed.

The gauge distribution $\mathfrak{X}(\mathcal{P}_\lambda) \cap \mathfrak{X}(\mathcal{P}_\lambda)^\perp$ is globally generated by $X_{\mathfrak{H}}$. The gauge freedom of the relativistic free particle is reflected in (6E.20) by the presence of the arbitrary multiplier k, and obviously corresponds to time reparametrizations. When $\zeta^A = 0$ the evolution is purely gauge, as predicted by Remark 6E.18.

b Electromagnetism. Since $\mathfrak{E}^0 = 0$ is the only primary constraint in Maxwell's theory on a fixed background spacetime, the polar $\mathfrak{X}(\mathcal{P}_\lambda)^\perp$ is spanned by $\delta/\delta A_0$. From expression (6C.23) for the electromagnetic Hamiltonian, we compute that $\delta H_{\lambda,\zeta}/\delta A_0 = 0$ iff

$$D_i \mathfrak{E}^i = 0, \tag{6E.21}$$

where we have performed an integration by parts. This is Gauss' Law, and defines $\mathcal{P}^2_{\lambda,\zeta} \subset \mathcal{P}_\lambda$. Continuing with the constraint algorithm, observe that along with $\delta/\delta A_0$, $\mathfrak{X}(\mathcal{P}^2_{\lambda,\zeta})^\perp$ is generated by vector fields of the form $V = (D_i a)\delta/\delta A_i$, where $a : \mathcal{P}^2_{\lambda,\zeta} \to \mathcal{F}(\Sigma_\lambda)$ is arbitrary (cf. (5A.6)). But then a computation gives

$$V[H_{\lambda,\zeta}] = \int_{\Sigma_\lambda} (\zeta^j a_{,j})_{,i} \mathfrak{E}^i \, d^3x_0 = -\int_{\Sigma_\lambda} \zeta^j a_{,j} \mathfrak{E}^i{}_{,i} \, d^3x_0$$

which vanishes by virtue of (6E.21). Thus the algorithm terminates with $\mathcal{C}_\lambda = \mathcal{P}^2_{\lambda,\zeta}$. Note that $\mathcal{C}_\lambda$ is indeed independent of the choice of slicing generator ζ, as promised by Corollary 6E.12. Moreover, it is obvious from (6E.21) that $\mathfrak{X}(\mathcal{C}_\lambda)^\perp \subset \mathfrak{X}(\mathcal{C}_\lambda)$ so $\mathcal{C}_\lambda$ is coisotropic and, in fact, all constraints are first class.

Maxwell's equations in the canonical form (6E.16) are satisfied by the vector field

$$X_\lambda = \left(\frac{dA_0}{d\lambda}\right)\frac{\delta}{\delta A_0} + \left(\frac{dA_i}{d\lambda}\right)\frac{\delta}{\delta A_i} + \left(\frac{d\mathfrak{E}^i}{d\lambda}\right)\frac{\delta}{\delta \mathfrak{E}^i}$$

provided

$$\frac{dA_i}{d\lambda} = \zeta^0 N \gamma^{-1/2}\gamma_{ij}\mathfrak{E}^j + \frac{1}{N\sqrt{\gamma}}(\zeta^0 M^j + \zeta^j)\mathfrak{F}_{ji} + D_i(\zeta^\mu A_\mu - \chi) \tag{6E.22}$$

and

$$\frac{d\mathfrak{E}^i}{d\lambda} = D_j\Big(\zeta^0\gamma^{ik}\gamma^{jm}\mathfrak{F}_{km} + \big[(\zeta^0 M^i + \zeta^i)\mathfrak{E}^j - (\zeta^0 M^j + \zeta^j)\mathfrak{E}^i\big]\Big). \tag{6E.23}$$

Equation (6E.22) reproduces the definition (6C.19) of the electric field density, while (6E.23) captures the dynamical content of Maxwell's theory. Note that $dA_0/d\lambda$ is left undetermined, in accord with the fact that $\delta/\delta A_0$ is a kinematic direction.

It is well-known that the canonical form (6E.22)–(6E.23) of Maxwell equations form a symmetric hyperbolic system when reduced to first order (McOwen [2003], Grundlach and Martín-García [2004]) and hence is well-posed provided Σ_λ is spacelike. (The computation is similar to that for bosonic strings, which we will explicitly carry out in Example **d** following.) One can also verify hyperbolicity on the covariant level as follows. Since the 4-dimensional form of the Maxwell equations in the Lorentz gauge $A^\mu{}_{;\mu} = 0$ reduce to wave equations for the A^ν (and hence are hyperbolic), and the gauge itself satisfies the wave equation, this theory is again seen to be well-posed provided Σ_λ is spacelike.[18] See Misner et al. [1973] and Wald [1984] for details here.

On a Minkowskian background relative to the slicing (6C.24), (6E.22) and (6E.23) take their more familiar forms

$$\frac{dA_i}{d\lambda} = \mathfrak{E}_i + D_i(A_0 - \chi) \tag{6E.24}$$

and

$$\frac{d\mathfrak{E}^i}{d\lambda} = D_j\mathfrak{F}^{ij}. \tag{6E.25}$$

Of course, X_λ given by (6E.22) and (6E.23) is not uniquely fixed; one can add to it any vector field $V \in \mathfrak{X}(\mathcal{C}_\lambda)^\perp$. Such a V has the form

$$V = a_0\frac{\delta}{\delta A_0} + D_i a\frac{\delta}{\delta A_i}$$

for arbitrary maps $a_0, a : \mathcal{C}_\lambda \to \mathcal{F}(\Sigma_\lambda)$. The first term in V simply reiterates the fact that the evolution of A_0 is arbitrary. To understand the significance

[18] In fact, to check well-posedness of a theory with gauge freedom in a spacetime with closed Cauchy surfaces, it is enough to verify this property in a particular gauge (Choquet-Bruhat [2004], Klainerman and Nicoló [1999]).

of the second term in V, it is convenient to perform a transverse-longitudinal decomposition of the spatial 1-form $\mathbf{A} = i_\lambda^* A$. (For simplicity, we return to the case of a Minkowskian background with the slicing (6C.24).) So split $\mathbf{A} = \mathbf{A}_T + \mathbf{A}_L$, where $\mathbf{A}_T$ is divergence-free and $\mathbf{A}_L$ is exact. Then (6E.24) splits into two equations:

$$\frac{d\mathbf{A}_T}{d\lambda} = \mathfrak{E} \quad \text{and} \quad \frac{d\mathbf{A}_L}{d\lambda} = \nabla A_0 - \nabla \chi.$$

(Note that the electric field is transverse by virtue of (6E.21).) The effect of the second term in V is to thus make the evolution of the longitudinal piece $\mathbf{A}_L$ completely arbitrary. In summary, both the temporal and longitudinal components A_0 and $\mathbf{A}_L$ of the potential A are gauge degrees of freedom whose conjugate momenta are constrained to vanish, leaving the transverse part $\mathbf{A}_T$ of A and its conjugate momentum $\mathfrak{E}$ as the true dynamical variables of the electromagnetic field.

As a corollary of this analysis, we observe that electromagnetism on a (1+1)-dimensional background is purely gauge.

The parametrized theory works much the same way. The main difference is that Hamilton's equations (6E.3) and the constraint conditions (6E.4) must now be understood as holding when evaluated in "variational directions" only (cf. Example **b** in §3B). By virtue of the metric primary constraints (6C.26), from (5C.9) we compute that $\mathfrak{X}(\tilde{\mathcal{P}}_\lambda)^\perp$ is spanned by the $\delta/\delta g_{\sigma\rho}$ in addition to $\delta/\delta A_0$. But, being nonvariational directions, the $\delta/\delta g_{\sigma\rho}$ cannot be used to generate secondary constraints,[19] so the constraint algorithm proceeds just as before. Similarly, Hamilton's equations now take the form

$$V \mathbin{\lrcorner} (\mathbf{i}_{X_\lambda}\tilde{\omega}_\lambda - \mathbf{d}H_{\lambda,\zeta}) \,|\, \tilde{\mathcal{P}}^2_{\lambda,\zeta} = 0$$

where X_λ and V are purely variational vector fields, i.e., can have no components in the $\delta/\delta g_{\sigma\rho}$ directions. Consequently, these equations are exactly equivalent to (6E.22) and (6E.23); in particular, there are no equations of motion for the parameters $g_{\sigma\rho}$. Note also that $H_{\lambda,\zeta} \mid \tilde{\mathcal{P}}^2_{\lambda,\zeta} \neq 0$ even though the theory is parametrized; the reason is that the metric g is not variational.

[19] If one attempted to use these directions in the constraint algorithm, one would be led to insist that

$$\frac{\delta H_{\lambda,\zeta}}{\delta g_{\sigma\rho}} = -\frac{1}{2}\zeta^0 \mathfrak{T}^{\sigma\rho} = 0,$$

where $\mathfrak{T}^{\sigma\rho}$ is the stress-energy-momentum tensor of the electromagnetic field—a clearly inappropriate requirement.

c A Topological Field Theory. From (6C.30) we have the instantaneous primary constraint set

$$\mathcal{P}_\lambda = \left\{(A, \pi) \in T^*\mathcal{Y}_\lambda \mid \pi^0 = 0 \;\text{ and }\; \pi^i = \epsilon^{0ij} A_j\right\}.$$

It follows that $\mathfrak{X}(\mathcal{P}_\lambda)^\perp$ is spanned by the vector field $\delta/\delta A_0$. With the Hamiltonian $H_{\lambda,\zeta}$ given by (6C.32), insisting that $\delta H_{\lambda,\zeta}/\delta A_0 = 0$ produces the spatial flatness condition (recall that $n = 2$)

$$F_{12} = 0. \tag{6E.26}$$

This equation defines $\mathcal{P}^2_{\lambda,\zeta} \subset \mathcal{P}_\lambda$. Proceeding, we note that along with $\delta/\delta A_0$, $\mathfrak{X}(\mathcal{P}^2_{\lambda,\zeta})^\perp$ is generated by vector fields of the form

$$V = D_i a \left(\epsilon^{0ij} \frac{\delta}{\delta \pi^j} - \frac{\delta}{\delta A_i}\right),$$

where $a : \mathcal{P}^2_{\lambda,\zeta} \to \mathcal{F}(\Sigma_\lambda)$ is arbitrary. But then a computation gives

$$V[H_{\lambda,\zeta}] = \frac{1}{2} \int_{\Sigma_\lambda} \epsilon^{0ij} \zeta^m a_{,m} F_{ij}\, d^3x_0$$

which vanishes in view of (6E.26). Thus the constraint algorithm terminates with $\mathcal{C}_\lambda = \mathcal{P}^2_{\lambda,\zeta}$.

Since $\mathfrak{X}(\mathcal{C}_\lambda)^\perp \subset \mathfrak{X}(\mathcal{C}_\lambda)$, $\mathcal{C}_\lambda$ is coisotropic in $\mathcal{P}_\lambda$, whence the secondary constraint (6E.26) is first class. The primary constraint $\pi^0 = 0$ is also first class, while the remaining two primaries $\pi^i - \epsilon^{0ij} A_j = 0$ are second class.

Next, suppose the vector field

$$X_\lambda = \left(\frac{dA_0}{d\lambda}\right) \frac{\delta}{\delta A_0} + \left(\frac{dA_i}{d\lambda}\right) \frac{\delta}{\delta A_i} + \left(\frac{d\pi^i}{d\lambda}\right) \frac{\delta}{\delta \pi^i}$$

satisfies the Chern–Simons equations in the Hamiltonian form (6E.16). Then by (6C.31) we must have

$$\frac{dA_i}{d\lambda} = D_i(\zeta^\mu A_\mu), \tag{6E.27}$$

and from (6C.29) we then derive

$$\frac{d\pi^i}{d\lambda} = \epsilon^{0ij} D_j(\zeta^\mu A_\mu). \tag{6E.28}$$

As in electromagnetism, $\delta/\delta A_0$ is a kinematic direction with the consequence that $dA_0/d\lambda$ is left undetermined. By subtracting $dA_i/d\lambda$ given by (6E.27) from

$$\dot{A}_i = \zeta^\mu D_\mu A_i + A_\mu \zeta^\mu{}_{,i}$$

obtained from (6B.1) while taking (6C.27) into account, we get

$$\zeta^\mu(D_i A_\mu - D_\mu A_i) = 0$$

which, when combined with (6E.26), yields the remaining flatness conditions $F_{i0} = 0$ in (3B.23). Equation (6E.28) yields nothing new.

Finally, note that (i) when restricted to $\mathcal{C}_\lambda$ the Chern–Simons Hamiltonian (6C.32) vanishes by (6E.26), and (ii) we may rearrange

$$X_\lambda = \left(\frac{dA_0}{d\lambda}\right)\frac{\delta}{\delta A_0} - D_i(\zeta^\mu A_\mu)\left(\epsilon^{0ij}\frac{\delta}{\delta\pi^j} - \frac{\delta}{\delta A_i}\right) \in \mathfrak{X}(\mathcal{C}_\lambda)^\perp,$$

so that the Chern–Simons evolution is completely gauge, as must be the case for a parametrized field theory in which all fields are variational.

One way to see that the Chern–Simons equations $F_{\mu\nu} = 0$ make up a well-posed system is to observe that if we make the gauge choices $A_0 = 0$ and $\zeta_X = (1, \mathbf{0})$, then the field equations imply that $\partial_0 A_\nu = 0$, which clearly determines a unique solution given initial data consisting of A_i satisfying $A_{[1,2]} = 0$.

d Bosonic Strings. From (6C.36) and (6C.38) we see that $\mathfrak{X}(\mathcal{P}_\lambda)^\perp$ is spanned by the vector fields $\delta/\delta h_{\sigma\rho}$ or, equivalently, $\delta/\delta h^{\sigma\rho}$. Now demand that $\delta H_{\lambda,\zeta}/\delta h^{\sigma\rho} = 0$, where $H_{\lambda,\zeta}$ is given by (6C.37). For $(\sigma,\rho) = (1,1)$, this yields

$$\mathfrak{H} = \frac{1}{2\sqrt{\gamma}}\left(\pi^2 + D\varphi^2\right) = 0. \tag{6E.29}$$

Substituting this back into the Hamiltonian and setting $(\sigma,\rho) = (0,1)$, we get

$$\mathfrak{J} = \pi \cdot D\varphi = 0. \tag{6E.30}$$

Setting $(\sigma,\rho) = (0,0)$ produces nothing new, so that (6E.29) and (6E.30) are the only secondary constraints. Note that together they imply $H_{\lambda,\zeta}\,\big|\,\mathcal{P}^2_{\lambda,\zeta} = 0$, which of course reflects the fact that the bosonic string is a parametrized theory (and also that the slicing is a gauge slicing). As the notation suggests, $\mathfrak{H}$ and $\mathfrak{J}$ are the analogues, for bosonic strings, of the superhamiltonian and supermomentum, respectively, in ADM gravity.

For $N, M \in \mathcal{F}(\Sigma_\lambda)$, consider the Hamiltonian vector fields

$$\begin{aligned} X_{N\mathfrak{H}} &= \frac{N}{\sqrt{\gamma}} g^{AB}\pi_B \frac{\delta}{\delta\varphi^A} + \frac{1}{\sqrt{\gamma}} g_{AB} D(ND\varphi^B)\frac{\delta}{\delta\pi_A} \\ X_{M\mathfrak{J}} &= MD\varphi^A \frac{\delta}{\delta\varphi^A} + D(M\pi_A)\frac{\delta}{\delta\pi_A} \end{aligned} \tag{6E.31}$$

of $N\mathfrak{H}$ and $M\mathfrak{J}$, respectively. One verifies that $X_{N\mathfrak{H}}$ and $X_{M\mathfrak{J}}$, together with the $\delta/\delta h_{\sigma\rho}$, generate $\mathfrak{X}(\mathcal{P}^2_{\lambda,\zeta})^\perp = \mathfrak{X}(\mathcal{P}^2_{\lambda,\zeta})^\perp \cap \mathfrak{X}(\mathcal{P}^2_{\lambda,\zeta}) \subset \mathfrak{X}(\mathcal{P}^2_{\lambda,\zeta})$. Since in addition the Hamiltonian vanishes on $\mathcal{P}^2_{\lambda,\zeta}$, it follows that the constraint algorithm stops with $\mathcal{P}^2_{\lambda,\zeta} = \mathcal{C}_\lambda$ and also that all constraints are first class.

Writing the evolution vector field as

$$X_\lambda = \left(\frac{d\varphi^A}{d\lambda}\right)\frac{\delta}{\delta\varphi^A} + \left(\frac{d\pi_A}{d\lambda}\right)\frac{\delta}{\delta\pi_A} + \left(\frac{dh_{\sigma\rho}}{d\lambda}\right)\frac{\delta}{\delta h_{\sigma\rho}},$$

Hamilton's equations (6E.16) for the bosonic string are

$$\frac{d\varphi^A}{d\lambda} = -\frac{\zeta^0 N}{\sqrt{\gamma}} g^{AB}\pi_B - (\zeta^0 M + \zeta^1) D\varphi^A \tag{6E.32}$$

$$\frac{d\pi_A}{d\lambda} = -g_{AB} D\left(\frac{\zeta^0 N}{\sqrt{\gamma}} D\varphi^B\right) - D\left((\zeta^0 M + \zeta^1)\pi_A\right). \tag{6E.33}$$

Here the $dh_{\sigma\rho}/d\lambda$ are undetermined, which is a consequence of the fact that the $h_{\sigma\rho}$ are canonically conjugate to the first class primary constraints $\varpi^{\sigma\rho} = 0$, and hence are kinematic fields.

To establish well-posedness, we introduce $\vartheta^A = D\varphi^A$ and reduce (6E.32) and (6E.32) to first order, obtaining

$$\frac{d\vartheta^A}{d\lambda} = -(\zeta^0 M + \zeta^1) D\vartheta^A - \frac{\zeta^0 N}{\sqrt{\gamma}} g^{AB} D\pi_B + \cdots$$

$$\frac{d\varphi^A}{d\lambda} = 0 + \cdots$$

$$\frac{d\pi_A}{d\lambda} = -g_{AB}\frac{\zeta^0 N}{\sqrt{\gamma}} D\vartheta^B - (\zeta^0 M + \zeta^1) D\pi_A + \cdots .$$

where the ellipsis denotes terms of zeroth order. This system is evidently symmetric hyperbolic whence the the evolution equations are well-posed relative to any Lagrangian slicing in which Σ_λ is spacelike with respect to $h_{\mu\nu}$.[20,21]

Since $X_\lambda \in \mathfrak{X}(\mathcal{C}_\lambda) \cap \mathfrak{X}(\mathcal{C}_\lambda)^\perp$ the evolution is totally gauge. The gauge transformations on the fields (φ^A, π_A) generated by the vector fields $X_{N\mathfrak{H}}$ and $X_{M\mathfrak{J}}$ for N, M arbitrary express the covariance of the bosonic string under diffeomorphisms of X. The complete indeterminacy of the metric h generated by the vector fields $\delta/\delta h_{\sigma\rho}$ is also a result of invariance under diffeomorphisms—which

[20] One can also see this by noting that (3B.31) is a wave equation.
[21] See Remark 6A.4.

in two dimensions implies that the conformal factor is the only possible degree of freedom in h, cf. Example **d** in §3B—coupled with conformal invariance—which implies that even this degree of freedom is gauge.

In our examples, we have encountered at most secondary constraints, and in Example **a** there were only primary constraints. This is typical: in mechanics it is rare to find (uncontrived) systems with secondary constraints, and in field theories at most secondary constraints are the rule. (Two exceptional cases are Palatini gravity, which has tertiary constraints (see Part V), and the KdV equation, which has only primary constraints (see Gotay [1988].) In principle, however, the constraint chain 6E.12) can have arbitrary length, but this has no physical significance.

7 The Energy-Momentum Map

In Chapter 4 we defined a covariant momentum mapping for a group $\mathcal{G}$ of covariant canonical transformations of the multisymplectic manifold Z. This chapter correlates those ideas with momentum mappings (in the usual sense) on the presymplectic manifold $\mathcal{Z}_\tau$ and the symplectic manifold $T^*\mathcal{Y}_\tau$, and introduces the energy-momentum map on $\mathcal{Z}_\tau$. We then show that this energy-momentum map projects to a function $\mathfrak{E}_\tau$ on the τ-primary constraint set $\mathcal{P}_\tau$, and that under certain circumstances, $\mathfrak{E}_\tau$ is identifiable with the negative of the Hamiltonian. This is the key result which enables us in Part III to prove that the final constraint set for first class theories coincides with $\mathfrak{E}_\tau^{-1}(0)$, when $\mathcal{G}$ is the gauge group of the theory.

7A Induced Actions on Fields

We first show how group actions on Y and Z, etc., can be extended to actions on fields. Given a left action of a group $\mathcal{G}$ on a bundle $\pi_{XK} : K \to X$ covering an action of $\mathcal{G}$ on X, we get an induced left action of $\mathcal{G}$ on the space $\mathcal{K}$ of sections of π_{XK} defined by

$$\eta_{\mathcal{K}}(\sigma) = \eta_K \circ \sigma \circ \eta_X^{-1} \tag{7A.1}$$

for $\eta \in \mathcal{G}$ and $\sigma \in \mathcal{K}$, which generalizes the usual push-forward operation on tensor fields. The infinitesimal generator $\xi_{\mathcal{K}}(\sigma)$ of this action is simply the (negative of the) Lie derivative:

$$\xi_{\mathcal{K}}(\sigma) = -\pounds_\xi \sigma = \xi_K \circ \sigma - T\sigma \circ \xi_X. \tag{7A.2}$$

We consider the relationship between transformations of the spaces Z, $\mathcal{Z}$, and $\mathcal{Z}_\tau$. Let $\eta_Z : Z \to Z$ be a covariant canonical transformation covering $\eta_X : X \to X$ with the induced transformation $\eta_{\mathcal{Z}} : \mathcal{Z} \to \mathcal{Z}$ on fields given by (7A.1). For each $\tau \in \mathrm{Emb}(\Sigma, X)$, $\eta_{\mathcal{Z}}$ restricts to the mapping

$$\eta_{\mathcal{Z}_\tau} : \mathcal{Z}_\tau \to \mathcal{Z}_{\eta_X \circ \tau}$$

defined by

$$\eta_{\mathcal{Z}_\tau}(\sigma) = \eta_Z \circ \sigma \circ \eta_\tau^{-1}, \tag{7A.3}$$

where $\eta_\tau := \eta_X \,\big|\, \Sigma_\tau$ is the induced diffeomorphism from Σ_τ to $\eta_X(\Sigma_\tau)$.

Proposition 7A.1. *$\eta_{\mathcal{Z}_\tau}$ is a canonical transformation relative to the two-forms Ω_τ and $\Omega_{\eta_X \circ \tau}$; that is,*

$$(\eta_{\mathcal{Z}_\tau})^* \Omega_{\eta_X \circ \tau} = \Omega_\tau.$$

Proof. From equation (7A.3)

$$T\eta_{\mathcal{Z}_\tau} \cdot V = T\eta_Z \circ (V \circ \eta_\tau^{-1}) \tag{7A.4}$$

for $V \in T_\sigma \mathcal{Z}_\tau$. Thus,

$$
\begin{aligned}
&(\eta_{\mathcal{Z}_\tau})^* \Omega_{\eta_X \circ \tau}(V, W) \\
&\quad = \Omega_{\eta_X \circ \tau} \left(T\eta_Z \cdot V \circ \eta_\tau^{-1}, T\eta_Z \cdot W \circ \eta_\tau^{-1} \right) && \text{(by (7A.4))} \\
&\quad = \int_{\eta_X(\Sigma_\tau)} (\eta_Z \circ \sigma \circ \eta_\tau^{-1})^* (\mathbf{i}_{T\eta_Z \cdot W \circ \eta_\tau^{-1}} \mathbf{i}_{T\eta_Z \cdot V \circ \eta_\tau^{-1}} \Omega) && \text{(by (5B.8))} \\
&\quad = \int_{\eta_X(\Sigma_\tau)} (\eta_\tau^{-1})^* [\sigma^* \eta_Z^* (\mathbf{i}_{T\eta_Z \cdot W} \mathbf{i}_{T\eta_Z \cdot V} \Omega)] \\
&\quad = \int_{\Sigma_\tau} (\sigma^* \eta_Z{}^*)(\mathbf{i}_{T\eta_Z \cdot W} \mathbf{i}_{T\eta_Z \cdot V} \Omega) && \text{(change of variables formula)} \\
&\quad = \int_{\Sigma_\tau} \sigma^* (\mathbf{i}_W \mathbf{i}_V \eta_Z{}^* \Omega) && \text{(by naturality of pull-back)} \\
&\quad = \int_{\Sigma_\tau} \sigma^* (\mathbf{i}_W \mathbf{i}_V \Omega) && \text{(since } \eta \text{ is covariant canonical)} \\
&\quad = \Omega_\tau(V, W). && \text{(by (5B.8))}
\end{aligned}
$$

■

Similarly, one shows the following:

Proposition 7A.2. *If $\eta_{\mathcal{Z}} : \mathcal{Z} \to \mathcal{Z}$ is a special covariant canonical transformation, then $\eta_{\mathcal{Z}_\tau}$ is a special canonical transformation.*

7B The Energy-Momentum Map

Let $\mathcal{G}$ be a group acting by covariant canonical transformations on Z and let

$$J : Z \to \mathfrak{g}^* \otimes \Lambda^n Z$$

be a corresponding covariant momentum mapping. This induces the map $E_\tau : \mathcal{Z}_\tau \to \mathfrak{g}^*$ defined by

$$\langle E_\tau(\sigma), \xi \rangle = \int_{\Sigma_\tau} \sigma^* \langle J, \xi \rangle \tag{7B.1}$$

where $\xi \in \mathfrak{g}$ and $\langle J, \xi \rangle : Z \to \Lambda^n Z$ is defined by $\langle J, \xi \rangle(z) := \langle J(z), \xi \rangle$. While E_τ is not a momentum map in the usual sense on $\mathcal{Z}_\tau$—since $\mathcal{G}$ does not necessarily act on $\mathcal{Z}_\tau$—it will be shown later to be closely related to the Hamiltonian in the instantaneous formulation of classical field theory. For this reason we shall call E_τ the ***energy-momentum map***. Further justification for this terminology is given in the interlude following this chapter.

For actions on Z lifted from actions on Y, using adapted coordinates and (4C.7), (7B.1) becomes

$$\begin{aligned}\langle E_\tau(\sigma), \xi \rangle &= \int_{\Sigma_\tau} \sigma^* \big((p_A{}^\mu \xi^A + p\, \xi^\mu)\, d^n x_\mu - p_A{}^\mu \xi^\nu dy^A \wedge d^{n-1} x_{\mu\nu} \big) \\ &= \int_{\Sigma_\tau} \big((p_A{}^0 \xi^A + p\, \xi^0)\, d^n x_0 - p_A{}^\mu \xi^\nu \sigma^A{}_{,i}\, \sigma^* (dx^i \wedge d^{n-1} x_{\mu\nu}) \big)\end{aligned}$$

where the integrands are regarded as functions of x^i and where we write, in coordinates, $\sigma(x^i) = (x^i, \sigma^A(x^i), p(x^i), p_A{}^\mu(x^i))$. Since

$$dx^i \wedge d^{n-1} x_{\mu\nu} = \delta^i_\nu\, d^n x_\mu - \delta^i_\mu\, d^n x_\nu,$$

the expression above can be written in the form

$$\langle E_\tau(\sigma), \xi \rangle = \int_{\Sigma_\tau} \big(p_A{}^0 (\xi^A - \xi^i \sigma^A{}_{,i}) + (p + p_A{}^i \sigma^A{}_{,i}) \xi^0 \big)\, d^n x_0 \tag{7B.2}$$

$$= \int_{\Sigma_\tau} \big(p_A{}^0 (\xi^A - \xi^\mu \sigma^A{}_{,\mu}) + (p + p_A{}^\mu \sigma^A{}_{,\mu}) \xi^0 \big)\, d^n x_0, \tag{7B.3}$$

where (7B.3) is obtained from (7B.2) by adding and subtracting the term $\xi^0 {p_A}^0 {\sigma^A}_{,0}$. (For this to make sense, we suppose that σ is the restriction to Σ_τ of a section of π_{XZ}. Of course, (7B.3) is independent of this choice of extension.)

To obtain a bona fide momentum map on $\mathcal{Z}_\tau$, we restrict attention to the subgroup $\mathcal{G}_\tau$ of $\mathcal{G}$ consisting of transformations which stabilize the image of τ; that is,

$$\mathcal{G}_\tau := \{\eta \in \mathcal{G} \mid \eta_X(\Sigma_\tau) = \Sigma_\tau\}. \tag{7B.4}$$

We emphasize that the condition $\eta_X(\Sigma_\tau) = \Sigma_\tau$ within (7B.4) does not mean that each point of Σ_τ is left fixed by η_X, but rather that η_X moves the whole Cauchy surface Σ_τ onto itself.

For any $\eta \in \mathcal{G}_\tau$, the map $\eta_\tau := \eta_X \big| \Sigma_\tau$ is an element of $\mathrm{Diff}(\Sigma_\tau)$. It follows from Proposition 7A.1 that

$$\eta_{\mathcal{Z}_\tau}(\sigma) = \eta_Z \circ \sigma \circ \eta_\tau^{-1} \tag{7B.5}$$

is a canonical action of $\mathcal{G}_\tau$ on $\mathcal{Z}_\tau$. From (7A.2), the infinitesimal generator of this action is

$$\xi_{\mathcal{Z}_\tau}(\sigma) = \xi_Z \circ \sigma - T\sigma \circ \xi_\tau, \tag{7B.6}$$

where ξ_τ generates η_τ.

Being a subgroup of $\mathcal{G}$, $\mathcal{G}_\tau$ has a covariant momentum map which is given by J followed by the projection from $\mathfrak{g}^* \otimes \Lambda^n Z$ to $\mathfrak{g}_\tau^* \otimes \Lambda^n Z$, where $\mathfrak{g}_\tau$ is the Lie algebra of $\mathcal{G}_\tau$. Note that in adapted coordinates, $\xi \in \mathfrak{g}_\tau$ when $\xi_X^0 = 0$ on Σ_τ. From (7B.1), the map J induces the map $J_\tau := E_\tau \big| \mathfrak{g}_\tau : \mathcal{Z}_\tau \to \mathfrak{g}_\tau^*$ given by

$$\langle J_\tau(\sigma), \xi\rangle = \int_{\Sigma_\tau} \sigma^*\langle J, \xi\rangle \tag{7B.7}$$

for $\xi \in \mathfrak{g}_\tau$.

Proposition 7B.1. *J_τ is a momentum map for the $\mathcal{G}_\tau$-action on $\mathcal{Z}_\tau$ defined by (7B.5), and it is* Ad^**-equivariant if J is.*

Proof. Let $V \in T_\sigma \mathcal{Z}\tau$ and let v be a π_{XZ}-vertical vector field on Z such that $V = v \circ \sigma$. If f_λ is the flow of v, let $\sigma_\lambda = f_\lambda \circ \sigma$ so that the curve $\sigma_\lambda \in \mathcal{Z}\tau$ has tangent vector V at $\lambda = 0$. Therefore, with J_τ defined by (7B.7),we have

$$\langle \mathbf{i}_V \mathbf{d} J_\tau(\sigma), \xi\rangle = \frac{d}{d\lambda}\left[\int_{\Sigma_\tau} \sigma_\lambda^*\langle J, \xi\rangle\right]\Bigg|_{\lambda=0} = \int_{\Sigma_\tau} \sigma^* \pounds_v \langle J, \xi\rangle.$$

But

$$\int_{\Sigma_\tau} \sigma^* \pounds_v \langle J, \xi\rangle = \int_{\Sigma_\tau} \sigma^*(\mathbf{d}\,\mathbf{i}_v \langle J, \xi\rangle + \mathbf{i}_v \mathbf{d}\langle J, \xi\rangle),$$

and since Σ is compact and boundaryless,

$$\int_{\Sigma_\tau} \sigma^*(\mathbf{d}\,\mathbf{i}_v \langle J, \xi\rangle) = \int_{\Sigma_\tau} \mathbf{d}\sigma^*(\mathbf{i}_v \langle J, \xi\rangle) = 0$$

by Stokes' theorem. Therefore, by the definition (4C.3) of a covariant momentum mapping,

$$\langle \mathbf{i}_V \mathbf{d} J_\tau(\sigma), \xi\rangle = \int_{\Sigma_\tau} \sigma^*(\mathbf{i}_v \mathbf{d}\langle J, \xi\rangle) = \int_{\Sigma_\tau} \sigma^*[\mathbf{i}_v \mathbf{i}_{\xi_Z} \Omega]. \tag{7B.8}$$

Note that ξ_Z need not be π_{XZ}-vertical, so we cannot yet use Lemma 5B.1.

Now for any $w \in T\Sigma_\tau$, we have

$$\sigma^*(\mathbf{i}_v \mathbf{i}_{T\sigma\cdot w}\Omega) = -\sigma^*(\mathbf{i}_{T\sigma\cdot w} \mathbf{i}_v \Omega) = -\mathbf{i}_W \sigma^*(\mathbf{i}_v \Omega) = 0$$

by the naturality of pull-back and the fact that $\sigma^*(\mathbf{i}_v \Omega)$ vanishes since it is an $(n+1)$-form on an n-manifold. In particular, for $w = \xi_\tau$, we have

$$\sigma^*(\mathbf{i}_v \mathbf{i}_{T\sigma\cdot\xi_\tau}\Omega) = 0.$$

Combining this result with (7B.8) and using the fact that $\xi_Z - T\sigma\cdot\xi_\tau$ is π_{XZ}-vertical, we get

$$\begin{aligned}\langle \mathbf{i}_V \mathbf{d} J_\tau(\sigma), \xi\rangle &= \int_{\Sigma_\tau} \sigma^*(\mathbf{i}_v \mathbf{i}_{\xi_Z - T\sigma\cdot\xi_\tau}\Omega) \\ &= \Omega_\tau(\xi_{\mathcal{Z}_\tau}, V)\end{aligned}$$

by (7B.6) and (5B.8). Thus J_τ is a momentum map.

To show that J_τ is Ad^*-equivariant, we verify that for $\eta \in \mathcal{G}_\tau$ and $\xi \in \mathfrak{g}_\tau$, J_τ satisfies the condition

$$\langle J_\tau(\sigma), \mathrm{Ad}_{\eta^{-1}}\,\xi\rangle = \langle J_\tau(\eta_{\mathcal{Z}_\tau}(\sigma)), \xi\rangle.$$

However, from (7B.7) and (4C.4), we have

$$\langle J_\tau(\sigma), \mathrm{Ad}_{\eta^{-1}}\,\xi\rangle = \int_{\Sigma_\tau} \sigma^* \langle J, \mathrm{Ad}_{\eta^{-1}}\,\xi\rangle = \int_{\Sigma_\tau} \sigma^* {\eta_Z}^* \langle J, \xi\rangle;$$

whereas from (7B.5), (7B.7), and the change of variables formula, we get

$$\begin{aligned}\langle J_\tau(\eta_{\mathcal{Z}_\tau}(\sigma)), \xi\rangle &= \int_{\Sigma_\tau} (\eta_Z \circ \sigma \circ \eta_\tau^{-1})^* \langle J, \xi\rangle \\ &= \int_{\Sigma_\tau} (\eta_\tau^{-1})^* \sigma^* {\eta_Z}^* \langle J, \xi\rangle \\ &= \int_{\Sigma_\tau} \sigma^* {\eta_Z}^* \langle J, \xi\rangle,\end{aligned}$$

thereby establishing the desired equality. ■

7C Induced Momentum Maps on $T^*\mathcal{Y}_\tau$

We now demonstrate how the group actions and momentum maps carry over from the multisymplectic context to the instantaneous formalism. Recall that the phase space $(T^*\mathcal{Y}_\tau, \omega_\tau)$ is the symplectic quotient of the presymplectic manifold $(\mathcal{Z}_\tau, \Omega_\tau)$ by the map R_τ. The key observation is that both the action of $\mathcal{G}_\tau$ and the momentum map J_τ pass to the quotient.

First consider a canonical transformation $\eta_{\mathcal{Z}_\tau} : \mathcal{Z}_\tau \to \mathcal{Z}_\tau$. Define a map $\eta_{T^*\mathcal{Y}_\tau} : T^*\mathcal{Y}_\tau \to T^*\mathcal{Y}_\tau$ as follows: For each $\pi \in T^*_\varphi\mathcal{Y}_\tau$, set

$$\eta_{T^*\mathcal{Y}_\tau}(\pi) = R_\tau(\eta_{\mathcal{Z}_\tau}(\sigma)) \tag{7C.1}$$

where σ is any element of $R_\tau^{-1}(\{\pi\})$.

Proposition 7C.1. *The map $\eta_{T^*\mathcal{Y}_\tau}$ is a canonical transformation.*

Proof. To begin, we must show that $\eta_{T^*\mathcal{Y}_\tau}$ is well-defined; that is

$$R_\tau(\eta_{\mathcal{Z}_\tau}(\sigma)) = R_\tau(\eta_{\mathcal{Z}_\tau}(\sigma')) \quad \text{whenever} \quad \sigma, \sigma' \in R_\tau^{-1}(\{\pi\}).$$

Since $\eta_{\mathcal{Z}_\tau}$ is a canonical transformation, it preserves the kernel of Ω_τ. But this kernel equals the kernel of TR_τ by Corollary 5C.3(*ii*). Therefore, $\eta_{\mathcal{Z}_\tau}$ preserves the fibers of R_τ, and so $\eta_{T^*\mathcal{Y}_\tau}$ is well defined.

Since $\eta_{\mathcal{Z}_\tau}$ is a diffeomorphism and R_τ is a submersion, $\eta_{T^*\mathcal{Y}_\tau}$ is a diffeomorphism. That the map $\eta_{T^*\mathcal{Y}_\tau}$ preserves the symplectic form ω_τ is a straightforward computation using (7C.1), Corollary 5C.3, and the definitions. ■

This proposition shows that the canonical action of $\mathcal{G}_\tau$ on $\mathcal{Z}_\tau$ gives rise to a canonical action of $\mathcal{G}_\tau$ on $T^*\mathcal{Y}_\tau$ such that R_τ is equivariant; that is, for $\eta \in \mathcal{G}_\tau$, the following diagram commutes:

$$\begin{tikzcd} \mathcal{Z}_\tau \arrow[r, "R_\tau"] \arrow[d, "\eta_{\mathcal{Z}_\tau}"'] & T^*\mathcal{Y}_\tau \arrow[d, "\eta_{T^*\mathcal{Y}_\tau}"] \\ \mathcal{Z}_\tau \arrow[r, "R_\tau"'] & T^*\mathcal{Y}_\tau \end{tikzcd}$$

Regarding momentum maps, we have:

Proposition 7C.2. *If J_τ is a momentum map for the action of $\mathcal{G}_\tau$ on $\mathcal{Z}_\tau$, then $\mathfrak{J}_\tau : T^*\mathcal{Y}_\tau \to \mathfrak{g}_\tau^*$ defined by the diagram*

$$\begin{tikzcd} \mathcal{Z}_\tau \arrow[d, "R_\tau"'] \arrow[dr, "J_\tau"] & \\ T^*\mathcal{Y}_\tau \arrow[r, "\mathfrak{J}_\tau"'] & \mathfrak{g}_\tau^* \end{tikzcd} \tag{7C.2}$$

is a momentum map for the induced action of $\mathcal{G}_\tau$ on $T^*\mathcal{Y}_\tau$. Further, if J_τ is Ad^**-equivariant, then so is $\mathfrak{J}_\tau$.*

Proof. This follows from the facts that R_τ is equivariant and $R_\tau{}^*\omega_\tau = \Omega_\tau$. ■

We emphasize that *the momentum map $\mathfrak{J}_\tau$*, which we have defined on $T^*\mathcal{Y}_\tau$, *corresponds to the action of $\mathcal{G}_\tau$ only*. For the full group $\mathcal{G}$, the corresponding energy-momentum map does *not* pass from $\mathcal{Z}_\tau$ to $T^*\mathcal{Y}_\tau$. However, as we will see in §7D, the energy-momentum map E_τ does project to the *primary constraint submanifold* in $T^*\mathcal{Y}_\tau$.

For lifted actions we are able to obtain explicit formulas for the energy-momentum and momentum maps on $\mathcal{Z}_\tau$ and $T^*\mathcal{Y}_\tau$ and the relationship between them. Suppose the action of $\mathcal{G}$ on Z is obtained by lifting an action of $\mathcal{G}$ on Y. Then $\eta \in \mathcal{G}$ maps $\mathcal{Y}_\tau$ to $\mathcal{Y}_{\eta_X \circ \tau}$ according to

$$\eta_{\mathcal{Y}_\tau}(\varphi) = \eta_Y \circ \varphi \circ \eta_\tau^{-1} \tag{7C.3}$$

where $\eta_\tau = \eta_X \,\big|\, \Sigma_\tau$. This in turn restricts to an action of $\mathcal{G}_\tau$ on $\mathcal{Y}_\tau$ given by the same formula, with the infinitesimal generator

$$\xi_{\mathcal{Y}_\tau}(\varphi) = \xi_Y \circ \varphi - T\varphi \circ \xi_\tau \tag{7C.4}$$

where $\xi_\tau = \xi_X \,\big|\, \Sigma_\tau$.

Corollary 7C.3. *For actions lifted from Y:*

(*i*) *The energy-momentum map on $\mathcal{Z}_\tau$ is*

$$\langle E_\tau(\sigma), \xi\rangle = \int_{\Sigma_\tau} \varphi^*(\mathbf{i}_{\xi_Y}\sigma) \tag{7C.5}$$

where $\xi \in \mathfrak{g}$, and $\varphi = \pi_{YZ} \circ \sigma$.

(*ii*) *The induced $\mathcal{G}_\tau$-action on $T^*\mathcal{Y}_\tau$ given by (7C.1) is the usual cotangent action; that is,*

$$\eta_{T^*\mathcal{Y}_\tau}(\pi) = (\eta_{\mathcal{Y}_\tau}^{-1})^*\pi.$$

(*iii*) *The corresponding induced momentum map $\mathcal{J}_\tau$ on $T^*\mathcal{Y}_\tau$ defined by (7C.2) is the standard one; that is,*

$$\langle \mathcal{J}_\tau(\varphi, \pi), \xi\rangle = \langle \pi, \xi_{\mathcal{Y}_\tau}(\varphi)\rangle = \int_{\Sigma_\tau} \pi\left(\xi_{\mathcal{Y}_\tau}(\varphi)\right) \tag{7C.6}$$

for $\xi \in \mathfrak{g}_\tau$. Moreover, the momentum maps J, J_τ, and $\mathcal{J}_\tau$ are all Ad^**-equivariant.*

Proof. To prove (*i*), substitute formula (4C.6) into (7B.1) and note that

$$\sigma^*\langle J, \xi\rangle = \sigma^*{\pi_{YZ}}^*\mathbf{i}_{\xi_Y}\sigma = \varphi^*\mathbf{i}_{\xi_Y}\sigma. \tag{7C.7}$$

To prove (*ii*) let $\eta \in \mathcal{G}_\tau$, $\pi = R_\tau(\sigma) \in T^*_\varphi\mathcal{Y}_\tau$ and $V \in T_{\eta_{\mathcal{Y}_\tau}(\varphi)}\mathcal{Y}_\tau$. Then

$$\begin{aligned}
&\langle \eta_{T^*\mathcal{Y}_\tau}(\pi), V\rangle \\
&= \langle R_\tau(\eta_{\mathcal{Z}_\tau}(\sigma)), V\rangle && \text{(by (7C.1))} \\
&= \int_{\Sigma_\tau} (\eta_{\mathcal{Y}_\tau}(\varphi))^*[\mathbf{i}_V(\eta_{\mathcal{Z}_\tau}(\sigma))] && \text{(by (5C.1))} \\
&= \int_{\Sigma_\tau} (\eta_\tau^{-1})^*\varphi^*{\eta_Y}^*[\mathbf{i}_V(\eta_{\mathcal{Z}_\tau}(\sigma))] && \text{(by (7C.3))} \\
&= \int_{\Sigma_\tau} \varphi^*{\eta_Y}^*[\mathbf{i}_V(\eta_{\mathcal{Z}_\tau}(\sigma))] && \text{(by the change of variables formula)} \\
&= \int_{\Sigma_\tau} \varphi^*[\mathbf{i}_{T\eta_Y^{-1}\cdot V}{\eta_Y}^*(\eta_{\mathcal{Z}_\tau}(\sigma))] \\
&= \int_{\Sigma_\tau} \varphi^*[\mathbf{i}_{T\eta_Y^{-1}\cdot V}\sigma] && \text{(by (4B.3))}
\end{aligned}$$

$$= \langle R_\tau(\sigma), T\eta_Y^{-1} \cdot V \rangle \qquad \text{(by (5C.1))}$$

$$= \langle \pi, T\eta_Y^{-1} \cdot V \rangle$$

$$= \langle (\eta_Y^{-1})^* \pi, V \rangle.$$

To prove (iii) we compute, taking into account (7C.2), (7C.5), and (7C.4),

$$\langle \mathcal{J}_\tau(R_\tau(\sigma)), \xi \rangle = \langle J_\tau(\sigma), \xi \rangle = \int_{\Sigma_\tau} \varphi^*(\mathbf{i}_{\xi_Y} \sigma)$$

$$= \int_{\Sigma_\tau} \varphi^*(\mathbf{i}_{\xi_Y - T\varphi \cdot \xi_\tau} \sigma) = \langle R_\tau(\sigma), \xi_{\mathcal{Y}_\tau}(\varphi) \rangle,$$

where we have used

$$\varphi^* \mathbf{i}_{T\varphi \cdot \xi_\tau} \sigma = \mathbf{i}_{\xi_\tau} \varphi^* \sigma = 0$$

since $\varphi^*\sigma$ is an $(n+1)$-form on the n-manifold Σ_τ.

Finally, equivariance follows from Propositions 4C.1, 7B.1, and 7C.2. ■

7D The Hamiltonian and the Energy-Momentum Map

In §7B we defined the energy-momentum map E_τ on $\mathcal{Z}_\tau$. Here we show that for lifted actions, E_τ projects to a well-defined function

$$\mathcal{E}_\tau : \mathcal{P}_\tau \to \mathfrak{g}^*$$

on the τ-primary constraint set, which we refer to as the "instantaneous energy-momentum map." This is the central object for our later analysis.

Let the group $\mathcal{G}$ act on Y and consider the lifted action of $\mathcal{G}$ on Z. Using (4C.5) rewrite formula (7B.1) as

$$\langle E_\tau(\sigma), \xi \rangle = \int_{\Sigma_\tau} \langle \mathfrak{E}_\tau(\sigma), \xi \rangle$$

for $\sigma \in \mathcal{Z}_\tau$ and $\xi \in \mathfrak{g}$, where

$$\langle \mathfrak{E}_\tau(\sigma), \xi \rangle = \sigma^*(\mathbf{i}_{\xi_Z} \Theta) \tag{7D.1}$$

defines the ***energy-momentum density*** $\mathfrak{E}_\tau$.

While $\mathfrak{E}_\tau$ does not directly factor through the reduction map to give an instantaneous energy-momentum density on $T^*\mathcal{Y}_\tau$, we nonetheless have:

Proposition 7D.1. *The energy-momentum density* $\mathfrak{E}_\tau$ *induces an instantaneous energy-momentum density on* $\mathcal{P}_\tau \subset T^*\mathcal{Y}_\tau$*.*

Proof. Given any $(\varphi, \pi) \in \mathcal{P}_\tau$, let σ be a holonomic lift of (φ, π) to $\mathcal{N}_\tau$ (cf. §6C). We claim that for any $x \in \Sigma_\tau$ and $\xi \in \mathfrak{g}$, the quantity

$$\langle \mathfrak{E}_\tau(\sigma)(x), \xi \rangle \in \Lambda^n_x \Sigma_\tau$$

depends only upon $j^1\varphi(x)$ and $\pi(x)$. Thus, setting

$$\langle \mathfrak{E}_\tau(\varphi, \pi)(x), \xi \rangle = \langle \mathfrak{E}_\tau(\sigma)(x), \xi \rangle \tag{7D.2}$$

defines the instantaneous energy-momentum density (which we denote by the same symbol $\mathfrak{E}_\tau$) on $\mathcal{P}_\tau$.

If $\xi_X(x)$ is transverse to Σ_τ, then (7D.1) combined with (6C.10) gives

$$\langle \mathfrak{E}_\tau(\sigma)(x), \xi \rangle = -\mathfrak{H}_{\tau,\xi}(\varphi, \pi)(x). \tag{7D.3}$$

On the other hand, if $\xi_X(x) \in T_x\Sigma_\tau$, then from (7C.7) we compute

$$\langle \mathfrak{E}_\tau(\sigma)(x), \xi \rangle = \varphi^*(\mathbf{i}_{\xi_Y(\varphi(x))}\sigma(x)) = \varphi^*(\mathbf{i}_{\xi_Y(\varphi(x)) - T_x\varphi \cdot \xi_X(x)}\sigma(x))$$

where we have used the same 'trick' as in the proof of Corollary 7C.3(*iii*). Since $\xi_Y - T\varphi \cdot \xi_X$ is π_{XY}-vertical, we can now apply (5C.2) to obtain

$$\begin{aligned} \langle \mathfrak{E}_\tau(\sigma)(x), \xi \rangle &= \big\langle R_\tau(\sigma)(x), \xi_Y(\varphi(x)) - T_x\varphi \cdot \xi_X(x) \big\rangle \\ &= \langle \pi(x), \xi_{\mathcal{Y}_\tau}(\varphi)(x) \rangle. \end{aligned} \tag{7D.4}$$

In either case, $\langle \mathfrak{E}_\tau(\sigma)(x), \xi \rangle$ depends only upon the values of φ, its first derivatives, and π along Σ_τ. Thus the definition (7D.2) is meaningful for any $\xi \in \mathfrak{g}$. ■

Integrating (7D.2), we get the ***instantaneous energy-momentum map*** $\mathcal{E}_\tau : \mathcal{P}_\tau \to \mathfrak{g}^*$ defined by

$$\langle \mathcal{E}_\tau(\sigma), \xi \rangle = \int_{\Sigma_\tau} \langle \mathfrak{E}_\tau(\varphi, \pi), \xi \rangle. \tag{7D.5}$$

Two cases warrant special attention:

Corollary 7D.2. *Let* $\xi \in \mathfrak{g}$*.*

(*i*) *If* ξ_X *is everywhere transverse to* Σ_τ*, then*

$$\langle \mathcal{E}_\tau(\varphi, \pi), \xi \rangle = -H_{\tau,\xi}(\varphi, \pi) \tag{7D.6}$$

(*ii*) *If* ξ_X *is everywhere tangent to* Σ_τ, *then*

$$\langle \mathcal{E}_\tau(\varphi, \pi), \xi \rangle = \langle \mathcal{J}_\tau(\varphi, \pi), \xi \rangle. \tag{7D.7}$$

Proof. Assertion (*i*) follows from (7D.3) and (*ii*) is a consequence of (7D.4) and (7C.6). ■

Remark 7D.3. In general, $\mathcal{E}_\tau$ is defined only on the primary constraint set $\mathcal{P}_\tau$, as $H_{\tau,\xi}$ is. However, if $\mathcal{G} = \mathcal{G}_\tau$, then $\mathcal{E}_\tau = \mathcal{J}_\tau$ is defined on all of $T^*\mathcal{Y}_\tau$. (It was not necessary that σ be a *holonomic* lift for the proof of the second part of Proposition 7D.1, corresponding to the case when $\xi_X(x) \in T_x\Sigma_\tau$.) ♦

Remark 7D.4. Although the instantaneous energy-momentum map can be identified with the Hamiltonian (when $\xi_X \pitchfork \Sigma_\tau$) and the momentum map $\mathcal{J}_\tau$ for $\mathcal{G}_\tau$ (when $\xi_X \parallel \Sigma_\tau$), it is important to realize that $\langle \mathcal{E}_\tau(\varphi, \pi), \xi \rangle$ is defined for any $\xi \in \mathfrak{g}$, regardless of whether or not it is everywhere transverse or tangent to Σ_τ. ♦

Remark 7D.5. The relation (7D.6) between the instantaneous energy-momentum map and the Hamiltonian is only asserted to be valid in the context of lifted actions; for more general actions, we do not claim such a relationship. Luckily, in most examples, lifted actions are the appropriate ones to consider. ♦

The instantaneous energy-momentum map $\mathcal{E}_\tau$ on $\mathcal{P}_\tau$ is the cornerstone of our work since, via (7D.6) above, it constitutes the fundamental link between dynamics and the gauge group. From it we will be able to correlate the notion of "gauge transformation" as arising from the gauge group action with that in the Dirac–Bergmann theory of constraints. This in turn will make it possible to "recover" the first class initial value constraints from $\mathcal{E}_\tau$ because, according to §6E, they are the generators of gauge transformations.

Remark 7D.6. Indeed, in Chapter 10 we will show that for parametrized theories in which all fields are variational, the final constraint set $\mathcal{C}_\tau \subset \mathcal{E}_\tau^{-1}(0)$. Combining this with the relation (7D.6), we see that for such theories the Hamiltonian (defined relative to a $\mathcal{G}$-slicing) must vanish "on shell;" that is, $H_{\tau,\xi} \,\big|\, \mathcal{C}_\tau = 0$ as predicted in Remark 6E.18. ♦

Thus, in some sense, the energy-momentum map encodes in a single geometric object virtually *all* of the physically relevant information about a given classical field theory: its dynamics, its initial value constraints and its gauge freedom. Momentarily, in Interlude II, we will see that $\mathcal{E}_\tau$ also incorporates the stress-energy-momentum tensor of a theory. It is these properties of $\mathcal{E}_\tau$ that will eventually enable us to achieve our main goal; viz., to write the evolution equations in adjoint form.

Examples

a Particle Mechanics. If $\mathcal{G} = \mathrm{Diff}(\mathbb{R})$ acts on $Y = \mathbb{R} \times Q$ by time reparametrizations, then from (4C.9) the energy-momentum map on $\mathcal{Z}_t = \mathbb{R} \times T^*Q$ is

$$\langle E_t(p, q^1, \cdots, q^N, p_1, \cdots, p_N), \chi\rangle = p\chi(t).$$

But $p = 0$ on $\mathcal{N}$ by virtue of the time reparametrization-covariance of $\mathcal{L}$, cf. example **a** in §4D. Thus the instantaneous energy momentum map on $\mathcal{P}_t = R_t(\mathcal{N}_t)$ vanishes. The subgroup $\mathcal{G}_t$ consists of those diffeomorphisms which fix $\tau(\Sigma) = t \in \mathbb{R}$. However, the actions of $\mathcal{G}_t$ on $\mathcal{Z}_t$ and on $T^*\mathcal{Y}_t = T^*Q$ are trivial.

If $\mathcal{G} = \mathrm{Diff}(\mathbb{R}) \times G$, where G acts only on the factor Q, then $G \subset \mathcal{G}_t$. In this case, $\mathcal{J}_t$ reduces to the usual momentum map on T^*Q.

b Electromagnetism. For electromagnetism on a fixed background with $\mathcal{G} = \mathcal{F}(X)$, we find from (4C.12) and (7B.1) that in adapted coordinates,

$$\langle E_\tau(A, p, \mathfrak{F}), \chi\rangle = \int_{\Sigma_\tau} \mathfrak{F}^{\nu 0}\chi_{,\nu}\, d^3x_0$$

for $\chi \in \mathcal{F}(X)$. Now $\mathcal{G} = \mathcal{G}_\tau$, so in this case $\mathcal{J}_\tau$ and $\mathcal{E}_\tau$ coincide. Using the expression above for E_τ, (7C.2), and $\mathfrak{E}^\nu = \mathfrak{F}^{\nu 0}$, we get

$$\langle \mathcal{J}_\tau(A, \mathfrak{E}), \chi\rangle = \int_{\Sigma_\tau} \mathfrak{E}^\nu \chi_{,\nu}\, d^3x_0 \tag{7D.8}$$

on $T^*\mathcal{Y}_\tau$. Note that this agrees with formula (7C.6). When restricted to the primary constraint set $\mathcal{P}_\tau \subset T^*\mathcal{Y}_\tau$ given by $\mathfrak{E}^0 = 0$, (7D.8) becomes

$$\langle \mathcal{E}_\tau(A, \mathfrak{E}), \chi\rangle = \int_{\Sigma_\tau} \mathfrak{E}^i \chi_{,i}\, d^3x_0. \tag{7D.9}$$

In the parametrized case, when $\mathcal{G} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$, E_τ is replaced by $\tilde{E}_\tau$ where, with the help of (4C.16),

$$\langle \tilde{E}_\tau(A, p, \mathfrak{F}; g, \rho), (\xi, \chi)\rangle$$

$$= \int_{\Sigma_\tau} \Big(\mathfrak{F}^{\nu 0}(-A_\mu \xi^\mu{}_{,\nu} - A_{\nu,\mu}\xi^\mu + \chi_{,\nu}) - \rho^{\sigma\rho 0}(2 g_{\sigma\rho}\xi^\nu{}_{,\rho} + g_{\sigma\rho,\mu}\xi^\mu)$$

$$+ \ (p + \mathfrak{F}^{\mu\nu} A_{\mu,\nu} + \rho^{\sigma\rho\mu} g_{\sigma\rho,\mu})\xi^0 \Big)\, d^3x_0. \tag{7D.10}$$

Since elements of $\mathcal{G}_\tau$ preserve Σ_τ, each $(\xi, \chi) \in \tilde{\mathfrak{g}}_\tau$ satisfies $\xi^0 \big| \Sigma_\tau = 0$. Then $\tilde{E}_\tau$ projects to the momentum map

$$\langle \tilde{\mathcal{J}}_\tau(A, \mathfrak{E}; g, \rho), (\xi, \chi)\rangle = \int_{\Sigma_\tau} \Big(\mathfrak{E}^\nu(-A_\mu \xi^\mu{}_{,\nu} - A_{\nu,i}\xi^i + \chi_{,\nu})$$

$$- \rho^{\sigma\rho 0}(2 g_{\sigma\rho}\xi^\nu{}_{,\rho} + g_{\sigma\rho,i}\xi^i) \Big)\, d^3x_0 \tag{7D.11}$$

for the action of $\tilde{\mathcal{G}}_\tau$ on $T^*\tilde{\mathcal{Y}}_\tau$.

On $\tilde{\mathcal{P}}_\tau$, $\tilde{E}_\tau$ induces the instantaneous energy-momentum map

$$\langle \tilde{\mathcal{E}}_\tau(A, \mathfrak{E}; g), (\xi, \chi)\rangle$$

$$= \int_{\Sigma_\tau} \left(\mathfrak{E}^i(-A_\mu \xi^\mu{}_{,i} - A_{i,\mu}\xi^\mu + \chi_{,i}) - \frac{1}{4}\mathfrak{F}^{\mu\nu} F_{\mu\nu}\xi^0 \right) d^3x_0,$$

where we have used (3B.14). Adding and subtracting $-\mathfrak{E}^i A_{\mu,i}\xi^\mu$ to the integrand and rearranging yields

$$\int_{\Sigma_\tau} \left(\mathfrak{E}^i(\chi - A_\mu\xi^\mu)_{,i} + \mathfrak{E}^i F_{ij}\xi^j + \left(\frac{1}{2}\mathfrak{E}^i F_{i0} - \frac{1}{4}\mathfrak{F}^{ij}F_{ij} \right)\xi^0 \right) d^3x_0.$$

Using (6C.17) and (6C.19) to express F_{i0} in terms of $\mathfrak{E}^i$ and $\mathfrak{F}_{ij}$, this eventually gives

$$\langle \tilde{\mathcal{E}}_\tau(A, \mathfrak{E}; g), (\xi, \chi)\rangle =$$

$$- \int_{\Sigma_\tau} \bigg[(\xi^\mu A_\mu - \chi)_{,i}\mathfrak{E}^i + \frac{1}{N\sqrt{\gamma}}(\xi^0 M^i + \xi^i)\mathfrak{E}^j \mathfrak{F}_{ij}$$

$$+ \xi^0 N \gamma^{-1/2}\Big(\frac{1}{2}\gamma_{ij}\mathfrak{E}^i\mathfrak{E}^j + \frac{1}{4N^2}\gamma^{ik}\gamma^{jm}\mathfrak{F}_{ij}\mathfrak{F}_{km} \Big) \bigg] d^3x_0 \tag{7D.12}$$

where we have again made use of the splitting (6B.8)–(6B.10) of the metric g.

c A Topological Field Theory. Since the Chern–Simons Lagrangian density is not equivariant with respect to the $\mathcal{G} = \mathrm{Diff}(X) \ltimes \mathcal{F}(X)$-action, we are not guaranteed that our theory as developed above will apply. So we must proceed by hand.

On $\mathcal{Z}_\tau$ the multimomentum map (4C.19) induces the map

$$\langle E_\tau(\sigma), (\xi, \chi)\rangle$$
$$= \int_{\Sigma_\tau} \left(p^{\nu 0}(-A_\mu {\xi^\mu}_{,\nu} - A_{\nu,\mu}\xi^\mu + \chi_{,\nu}) + (p + p^{\mu\nu} A_{\mu,\nu})\xi^0\right) d^2x_0.$$

Now E_τ projects to the genuine momentum map

$$\langle \mathcal{J}_\tau(A, \pi), (\xi, \chi)\rangle = \int_{\Sigma_\tau} \pi^\nu(-A_\mu {\xi^\mu}_{,\nu} - A_{\nu,i}\xi^i + \chi_{,\nu})\, d^2x_0 \tag{7D.13}$$

on $T^*\mathcal{Y}_\tau$. Similarly, from (3B.20), one verifies that E_τ projects to the "ersatz" instantaneous energy-momentum map

$$\langle \mathcal{E}_\tau(A), (\xi, \chi)\rangle$$
$$= \int_{\Sigma_\tau} \left(\epsilon^{0ij} A_j(-A_\mu {\xi^\mu}_{,i} - A_{i,\mu}\xi^\mu + \chi_{,i}) + \epsilon^{\mu\nu\rho} A_\rho A_{\nu,\mu}\xi^0\right) d^2x_0$$
$$= \int_{\Sigma_\tau} \epsilon^{0ij} \left(A_j(\chi - A_\mu\xi^\mu)_{,i} + A_j F_{ik}\xi^k + \frac{1}{2} A_0 F_{ij}\xi^0\right) d^2x_0 \tag{7D.14}$$

on $\mathcal{P}_\tau$.

Not surprisingly, $\langle \mathcal{E}_\tau, (\xi, \chi)\rangle$ fails to coincide with the Chern–Simons Hamiltonian (6C.32) (when ξ_X is transverse to Σ_τ) because of the term involving χ. Nonetheless, an integration by parts shows that they agree on the *final* constraint set, cf. (6E.26). Indeed, the extra term in $\mathcal{E}_\tau$ amounts to adding the first class constraint $F_{12} = 0$ to the Hamiltonian with Lagrange multiplier χ, and this is certainly permissible according to the discussion at the end of §6E. From a slightly different point of view, since the action of $\mathcal{F}(X)$ on J^1Y leaves the Lagrangian density invariant up to a divergence, its action on $T\mathcal{Y}_\tau$ will leave the instantaneous Lagrangian (6C.28) invariant. In fact, (7D.13) is just the momentum map for this action (compare (7D.8)).

Alternately, we could proceed by simply dropping the $\mathcal{F}(X)$-action. The above formulæ remain valid, provided the terms involving χ are removed. In this context (7D.14) will now of course be a genuine energy-momentum map.

d Bosonic Strings. For the bosonic string, (4C.25) eventually leads to the expression

$$\langle E_\tau(\sigma),(\xi,\lambda)\rangle = \int_{\Sigma_\tau} \big(-p_A{}^0\varphi^A{}_{,\mu}\xi^\mu + \rho^{\sigma\rho 0}(2\lambda h_{\sigma\rho} - h_{\sigma\nu}\xi^\nu{}_{,\rho} - h_{\rho\nu}\xi^\nu{}_{,\sigma} - h_{\sigma\rho,\nu}\xi^\nu) + (p + p_A{}^\mu\varphi^A{}_{,\mu} + \rho^{\sigma\rho\mu}h_{\sigma\rho,\mu})\xi^0\big)\, d^1x_0 \qquad (7D.15)$$

for the energy-momentum map on $\mathcal{Z}_\tau$.

Restricting to the subgroup $\mathcal{G}_\tau$, (7D.15) reduces to

$$\langle \mathcal{J}_\tau(\varphi,h,\pi,\varpi),(\xi,\lambda)\rangle = \int_{\Sigma_\tau} \big(-(\pi\cdot D\varphi)\xi^1 + 2\lambda\varpi^\sigma{}_\sigma\ -2\varpi^\sigma{}_\rho\xi^\rho{}_{,\sigma} - \varpi^{\sigma\rho}h_{\sigma\rho,1}\xi^1\big)\, d^1x_0 \quad (7D.16)$$

on $T^*\mathcal{Y}_\tau$, where we have used h to lower the index on ϖ.

Finally, making use of (3B.25)–((3B.27) and (6B.8)–(6B.10) in (7D.15), we compute on $\mathcal{P}_\tau$

$$\begin{aligned}\langle \mathcal{E}_\tau(\varphi,h,\pi),(\xi,\lambda)\rangle &= \int_{\Sigma_\tau}\left(\frac{1}{2h^{00}\sqrt{-h}}\,\xi^0(\pi^2 + D\varphi^2)\ + \left(\frac{h^{01}}{h^{00}}\xi^0 - \xi^1\right)(\pi\cdot D\varphi)\right) d^1x_0 \\ &= -\int_{\Sigma_\tau}\left(\frac{1}{2\sqrt{\gamma}}\xi^0 N(\pi^2 + D\varphi^2) + (\xi^0 M + \xi^1)(\pi\cdot D\varphi)\right) d^1x_0. \qquad (7D.17)\end{aligned}$$

Note that this expression does not depend on λ. When $\xi = (1,\mathbf{0})$, it reduces to

$$\langle \mathcal{E}_\tau(\varphi,h,\pi),((1,\mathbf{0}),\lambda)\rangle = -\int_{\Sigma_\tau}\left(\frac{1}{2\sqrt{\gamma}}N(\pi^2 + D\varphi^2) + M(\pi\cdot D\varphi)\right) d^1x_0$$

from which one can read off the string ***superhamiltonian***

$$\mathfrak{H} = \frac{1}{2\sqrt{\gamma}}(\pi^2 + D\varphi^2)$$

and the string ***supermomentum***

$$\mathfrak{J} = \pi\cdot D\varphi.$$

Thus as claimed in the Introduction we have $\mathcal{E} = -(\mathfrak{H},\mathfrak{J})$, that is, the superhamiltonian and supermomentum are the components of the instantaneous *energy-momentum* map. The supermomentum by itself is a component of the *momentum* map $\mathcal{J}_\tau$ for the group $\mathcal{G}_\tau$ which does act in the instantaneous formalism, unlike $\mathcal{G}$.

Interlude II—The Stress-Energy-Momentum Tensor[22]

For many years classical field theorists grappled with the problem of constructing a suitable stress-energy-momentum ("SEM") tensor for a given collection of fields. There are various candidates for this object; for instance, from spacetime translations via Noether's theorem one can build the so-called ***canonical*** SEM tensor density

$$\mathfrak{t}^{\mu}{}_{\nu} = L\delta^{\mu}{}_{\nu} - \frac{\partial L}{\partial v^{A}{}_{\mu}} v^{A}{}_{\nu}, \tag{II.1}$$

as is found in in, for example, Soper [1976], equation (3.3.3). Unfortunately, $\mathfrak{t}^{\mu\nu}$ is typically neither symmetric nor gauge-invariant, and much work has gone into efforts to "repair" it, cf. Belinfante [1939], Wentzel [1949], Corson [1953], and Davis [1970], General relativity provides an entirely different method of generating a SEM tensor (Hawking and Ellis [1973], Misner et al. [1973]). For a matter field coupled to gravity,

$$\mathfrak{T}^{\mu\nu} = 2\frac{\delta L}{\delta g_{\mu\nu}} \tag{II.2}$$

defines the ***Hilbert*** SEM tensor density. By its very definition, $\mathfrak{T}^{\mu\nu}$ is both symmetric and gauge-covariant. Despite its lack of an immediate physical interpretation, this "modern" construction of the SEM tensor has largely supplanted that based on Noether's theorem. Formulæ like this are also important in continuum mechanics, relativistic or not. For example, in nonrelativistic elasticity theory, (II.2) is sometimes called the ***Doyle–Eriksen formula*** and it defines the Cauchy stress tensor. This formula has been connected to the covariance of the theory by Marsden and Hughes [1983], and Simo and Marsden [1984]. For further development of this idea, see Yavari et al. [2006].

Discussions and other definitions of SEM tensors and related objects can be found in Souriau [1974], Kijowski and Tulczyjew [1979], as well as Ferraris and Francaviglia [1985, 1991], In general, however, the physical significance of these proposed SEM tensors remains unclear. In field theories on a Minkowski background, $\mathfrak{t}^{\mu\nu}$ is often symmetrized by adding to it a certain expression which is attributed to the energy density, momentum density, and stress arising from

[22] This is essentially a condensed version of Gotay and Marsden [1992].

spin. While one can give a definite prescription for carrying out this symmetrization Belinfante [1939], such modifications are nonetheless *ad hoc*. The situation is more problematical for field theories on a curved background, or for topological field theories in which there is no metric at all. Even for systems coupled to gravity, the definition (II.2) of $\mathcal{T}^{\mu\nu}$ has no *direct* physical significance. Now, the Hilbert tensor can be regarded as a constitutive tensor for the matter fields by virtue of the fact that it acts as the source of Einstein's equations. But this interpretation of $\mathcal{T}^{\mu\nu}$ is in a sense secondary, and it would be preferable to have its justification as a SEM tensor follow from first principles, i.e., from an analysis based on symmetries and Noether theory.

Moreover the relations between various SEM tensors, and in particular the canonical and Hilbert tensors, is somewhat obscure. On occasion, $\mathcal{T}^{\mu\nu}$ is obtained by directly symmetrizing $\mathfrak{t}^{\mu\nu}$ (as in the case of the Dirac field), but more often not (e.g., electromagnetism). For tensor or spinor field theories, Belinfante [1940] and Rosenfeld [1940] (see also Trautman [1965]) showed that $\mathcal{T}^{\mu}{}_{\nu}$ can be viewed as the result of "correcting" $\mathfrak{t}^{\mu}{}_{\nu}$:

$$\mathcal{T}^{\mu}{}_{\nu} = \mathfrak{t}^{\mu}{}_{\nu} + \nabla_{\rho} K_{\nu}^{\mu\rho} \tag{II.3}$$

for some quantities $K_{\nu}^{\mu\rho}$. We refer to (II.3) as the ***Belinfante–Rosenfeld formula***.

Our purpose in this Interlude is to show how the multisymplectic formalism we have developed can be used to give a *physically meaningful* definition of the SEM tensor based on covariance considerations for (essentially) arbitrary field theories that suffers none of these maladies. We will demonstrate that the SEM tensor density so defined (we call it $\mathfrak{T}^{\mu}{}_{\nu}$) satisfies a generalized version of the Belinfante–Rosenfeld formula (II.3), and hence *naturally* incorporates both $\mathfrak{t}^{\mu}{}_{\nu}$ and the "correction terms" which are necessary to make the latter gauge-covariant. Furthermore, in the presence of a metric on spacetime, we will show that our SEM tensor coincides with the Hilbert tensor, and hence is *automatically* symmetric.

As might be expected by now, the key ingredient in our analysis is the multimomentum map associated to the spacetime diffeomorphism group. We use it to define the SEM tensor density $\mathfrak{T}^{\mu}{}_{\nu}$ by means of fluxes across hypersurfaces in spacetime. This makes intuitive sense, since the multimomentum map describes how the fields "respond" to spacetime deformations. One main consequence is that *our definition uniquely determines* $\mathfrak{T}^{\mu}{}_{\nu}$; this is because our definition is "integral" (i.e., in terms of fluxes) as opposed to being based on differential conservation laws as is traditionally done, cf. Davis [1970]. Thus unlike, say, $\mathfrak{t}^{\mu}{}_{\nu}$, our SEM tensor is *not* merely defined up to a curl, and correspondingly there is no possibility of—and no necessity for—modifying it. The fact that the relevant group is the entire spacetime diffeomorphism group, and not just the translation group, is also crucial. Indeed, nonconstant deformations are what give rise to the "correction terms" mentioned above. Moreover, our analysis is then applicable to field theories on arbitrary spacetimes (in which context of course the translation group, let alone the Poincaré group, no longer has global meaning).

The SEM tensor measures the response of the fields to localized (i.e., compactly supported) spacetime deformations. We therefore assume that the Lagrangian field theory under consideration is parametrized, at least to the extent that that the image of the gauge group $\mathcal{G}$ under the natural projection $\mathrm{Aut}(Y) \to \mathrm{Diff}(X)$ contains the group $\mathrm{Diff}_c(X)$ of compactly supported diffeomorphisms. However, $\mathrm{Diff}_c(X)$ does not necessarily act on the fields, at least not *ab initio*. The reason is that $\mathrm{Diff}_c(X)$ is not naturally a subgroup of $\mathcal{G}$, but rather is a subgroup of the quotient group $\mathcal{G}/\mathcal{G}_{\mathrm{Id}}$, where $\mathcal{G}_{\mathrm{Id}}$ consists of those elements of $\mathcal{G}$ that cover the identity on X. Therefore, to define the SEM tensor, we need an embedding $\mathrm{Diff}_c(X) \to \mathcal{G}$, so that each element of $\mathrm{Diff}_c(X)$ gives rise to a gauge transformation. To be precise, assume there is a group monomorphism of $\mathrm{Diff}_c(X)$ into $\mathcal{G}$ such that the composition $\mathrm{Diff}_c(X) \to \mathcal{G} \to \mathrm{Diff}(X)$ is the identity. In general, $\mathcal{G}$ is larger than $\mathrm{Diff}_c(X)$, but the stress-energy-momentum tensor is associated only with the $\mathrm{Diff}_c(X)$ "part" of $\mathcal{G}$. For instance, in continuum mechanics in the inverse material representation, $\mathfrak{T}^{\mu\nu}$ is to measure the net energy flow, momentum flux and stress across hypersurfaces in spacetime—even if the material has internal structure.

The embedding $\mathrm{Diff}_c(X) \to \mathcal{G}$ of Lie groups determines, by differentiation, a Lie algebra monomorphism $\mathfrak{X}_c(X) \to \mathfrak{g}$ given by[23] $\xi \mapsto \xi_Y$, where $\xi^A = \xi^A(x^\mu, y^B, [\xi^\nu])$ is a smooth differential function of ξ^ν. We suppose that the action of $\mathrm{Diff}_c(X)$ on Y is "local" in the sense that ξ^A depends on ξ^ν and its derivatives up to some order $k < \infty$. We call k the ***differential index*** of the field theory; it is the order of the highest derivatives that appear in the transformation laws for the fields. The association $\xi \mapsto \xi_Y$ is linear in ξ^ν, so that in coordinates we may write

$$\xi^A = C^{A\rho_1\ldots\rho_k}_{\nu}\xi^\nu{}_{,\rho_1\ldots\rho_k} + \ldots + C^{A\rho}_{\nu}\xi^\nu{}_{,\rho} + C^A{}_\nu \xi^\nu,$$

[23] Here we are concretely viewing $\mathcal{G} \subset \mathrm{Aut}(Y)$, so that $\mathfrak{g} \subset \mathrm{aut}(Y)$.

where the coefficients $C^{A\rho_1\cdots\rho_r}_{\nu}$ for $r = 0, 1, \ldots, k$ depend only upon x^μ and y^B.

For tensor field theories, using the standard embedding $\eta \mapsto \eta_* = (\eta^{-1})^*$, one has simply

$$\xi^A = C^{A\rho}_{\nu}\xi^\nu{}_{,\rho},$$

so $k = 1$ in this instance (unless ϕ is a scalar field, in which case all the coefficient functions $C^{A\rho_1\cdots\rho_m}_{\nu}$ vanish). But if one uses a "nonstandard" embedding $\mathrm{Diff}_c(X) \to \mathcal{G}$, zeroth order terms may appear as well:

$$\xi^A = C^{A\rho}_{\nu}\xi^\nu{}_{,\rho} + C^A{}_\nu \xi^\nu. \tag{II.4}$$

We will illustrate this in the context of electromagnetism in the Examples following. For a field theory based on a linear connection on X (such as Palatini gravity), one would have $k = 2$. For the sake of simplicity, we henceforth suppose that $k \leq 1$ as is typically the case in applications, such as tensor and spinor field theories; the general situation is covered in Gotay and Marsden [1992].

Before proceeding to the definition of the SEM tensor, we derive the consequences of the covariance condition (4D.2). Substituting (II.4) into (4D.2) yields

$$\begin{aligned} L_{,\nu}\xi^\nu + L\xi^\nu{}_{,\nu} + L_A\left(C^{A\rho}_{\nu}\xi^\nu{}_{,\rho} + C^A{}_\nu\xi^\nu\right) \\ + L_A{}^\mu\left(C^{A\rho}_{\nu}\xi^\nu{}_{,\rho\mu} + (D_\mu C^{A\rho}_{\nu})\xi^\nu{}_{,\rho}\right. \\ \left. + C^A{}_\nu\xi^\nu{}_{,\mu} + (D_\mu C^A{}_\nu)\xi^\nu - v^A{}_\nu\xi^\nu{}_{,\mu}\right) = 0, \end{aligned}$$

where $D_\mu = \partial_\mu + v^A{}_\mu\partial_A$ is the is the formal or "total" derivative, and we have abbreviated $\partial L/\partial v^A{}_\mu$ by $L_A{}^\mu$, etc. Since $\mathrm{Diff}_c(X) \to \mathcal{G}$ is an *embedding*, ξ^ν and its derivatives are arbitrarily specifiable. Thus, equating to zero the coefficients of the $\xi^\nu{}_{,\rho_1\ldots\rho_m}$, we obtain the following results.

For $m = 2$:

$$C^{A(\rho_1}_{\nu} L_A{}^{\rho_2)} = 0.$$

For $m = 1$:

$$(L\delta^\rho{}_\nu - L_A{}^\rho v^A{}_\nu) + C^{A\rho}_{\nu}L_A + C^A{}_\nu L_A{}^\rho + (D_\mu C^{A\rho}_{\nu})L_A{}^\mu = 0.$$

For $m = 0$:

$$L_{,\nu} + C^A{}_\nu L_A + (D_\mu C^A{}_\nu)L_A{}^\mu = 0.$$

Using the Leibniz rule and setting

$$K_\nu^{\rho_1\cdots\rho_m\mu} := C^{A\rho_1\cdots\rho_m}_{\ \nu} L_A{}^\mu,$$

these can be rewritten

$$\left.\begin{aligned} K_\nu^{(\rho_1\rho_2)} &= 0 \\ (L\delta^\rho{}_\nu - L_A{}^\rho v^A{}_\nu) + K_\nu^\rho + D_\mu K_\nu^{\rho\mu} + C^{A\rho}_{\ \nu}\frac{\delta L}{\delta y^A} &= 0 \\ L_{,\nu} + D_\mu K_\nu^\mu + C^A{}_\nu \frac{\delta L}{\delta y^A} &= 0. \end{aligned}\right\} \tag{II.5}$$

We are now ready to state our main result, which characterizes fluxes of the multimomentum map $J^{\mathcal{L}}$ in the Lagrangian representation across hypersurfaces in X.

Theorem II.1. *Consider a $\mathcal{G}$-covariant Lagrangian field theory. For each section $\phi : X \to Y$ there exists a unique $(1,1)$-tensor density $\mathfrak{T}(\phi)$ on X such that*

$$\int_\Sigma i_\Sigma^*(j^1\phi)^* J^{\mathcal{L}}(\xi_Y) = \int_\Sigma \mathfrak{T}^\mu{}_\nu(\phi)\xi^\nu d^n x_\mu \tag{II.6}$$

for all $\xi \in \mathfrak{X}_c(X)$ and all hypersurfaces Σ, where $i_\Sigma : \Sigma \to X$ is the inclusion.[24]

We call $\mathfrak{T}(\phi)$ the ***SEM tensor density*** of the field ϕ.

Proof. Recall that we take $k \le 1$. Using (4D.8) and (II.4), the left hand side of (II.6) becomes

$$\begin{aligned} &\int_\Sigma i_\Sigma^*(j^1\phi)^* J^{\mathcal{L}}(\xi_Y) \\ &\qquad = \int_\Sigma \Big(L_A{}^\mu\big(\xi^A - v^A{}_\nu\xi^\nu\big) + L\xi^\mu\Big) d^n x_\mu \\ &\qquad = \int_\Sigma \Big(L_A{}^\mu\big(C^{A\rho}_{\ \nu}\xi^\nu{}_{,\rho} + C^A{}_\nu\xi^\nu - v^A{}_\nu\xi^\nu\big) + L\delta^\mu{}_\nu\xi^\nu\Big) d^n x_\mu \\ &\qquad = \int_\Sigma \Big(K_\nu^{\rho\mu}\xi^\nu{}_{,\rho} + \big(L\delta^\mu{}_\nu - L_A{}^\mu v^A{}_\nu + K_\nu^\mu\big)\xi^\nu\Big) d^n x_\mu. \end{aligned} \tag{II.7}$$

[24] Here Σ is completely arbitrary; it need not be noncompact nor without boundary. It also need not be spacelike in the presence of a metric on X.

Let $U \subset X$ be a chart in which $\Sigma \cap U$ is a hyperplane. By means of a partition of unity argument, it suffices to consider the case when the vector field ξ has support contained in U. Construct an $(n+1)$-dimensional region $V \subset X$ such that $\partial V = (\Sigma \cap U) \cup \Sigma'$, where $\xi \,|\, \Sigma' = 0$. By the divergence theorem, the first term in (II.7) becomes

$$\begin{aligned}\int_\Sigma K_\nu^{\rho\mu} \xi^\nu{}_{,\rho}\, d^n x_\mu &= \int_V \left(K_\nu^{\rho\mu} \xi^\nu{}_{,\rho}\right)_{,\mu} d^{n+1}x \\ &= \int_V \Big(K_\nu^{\rho\mu}{}_{,\mu} \xi^\nu{}_{,\rho} + K_\nu^{\rho\mu} \xi^\nu{}_{,\rho\mu}\Big) d^{n+1}x. \qquad \text{(II.8)}\end{aligned}$$

In (II.8) the second term on the last line vanishes by virtue of the first equation in (II.5). Applying the Leibniz rule to the first term yields

$$\int_V K_\nu^{\rho\mu}{}_{,\mu} \xi^\nu{}_{,\rho}\, d^{n+1}x = \int_V \Big(\left(K_\nu^{\rho\mu}{}_{,\mu} \xi^\nu\right)_{,\rho} - K_\nu^{\rho\mu}{}_{,\mu\rho} \xi^\nu\Big) d^{n+1}x.$$

Using the first equation of (II.5) once more, the second integrand here also vanishes by symmetry. Thus (II.8) reduces to

$$\int_V \left(K_\nu^{\rho\mu}{}_{,\mu} \xi^\nu\right)_{,\rho} d^{n+1}x = \int_\Sigma K_\nu^{\rho\mu}{}_{,\mu} \xi^\nu d^n x_\rho$$

again by the divergence theorem.

Substituting these results back into (II.7) and reindexing, we therefore obtain (II.6) with

$$\mathfrak{T}^\mu{}_\nu = L\delta^\mu{}_\nu - L_A{}^\mu v^A{}_\nu + K_\nu^\mu + D_\rho K_\nu^{\mu\rho}. \qquad \text{(II.9)}$$

Finally, it is clear from the fundamental lemma of the calculus of variations that $\mathfrak{T}^\mu{}_\nu$ so defined is unique. While we have derived the formula for the SEM tensor density in coordinates, in fact it is a tensor density as its defining property (II.6) is clearly intrinsic. ∎

Reverting to our original notation, this last formula becomes

$$\mathfrak{T}^\mu{}_\nu = L\delta^\mu{}_\nu - L_A{}^\mu v^A{}_\nu + L_A{}^\mu C^A{}_\nu + D_\rho(L_A{}^\rho C^A{}^\mu_\nu). \qquad \text{(II.10)}$$

Taking (II.1) into account, we see that

$$\mathfrak{T}^\mu{}_\nu = \mathfrak{t}^\mu{}_\nu + \text{ “correction terms”};$$

hence $\mathfrak{T}^\mu{}_\nu$ may be regarded as a modification of the canonical SEM tensor density. Formula (II.10) is thus a ***generalized Belinfante–Rosenfeld formula***.

Remark II.2. Formula (II.10) is consistent with (II.1) if one uses, instead of $\mathrm{Diff}_c(X)$, just translations in X and sets $\xi^A = 0$. Similarly, if one uses the Poincaré group in place of $\mathrm{Diff}_c(X)$ then, for tensor field theories, $\mathfrak{T}^\mu{}_\nu$ reduces to the canonical SEM tensor as modified by Belinfante [1939]; see also Wentzel [1949] and Corson [1953]. However, with our approach the "correction terms" in (II.10) *naturally* appear.

Remark II.3. Although we have derived (II.10) in the case when $k = 1$, it happens that *this expression is valid for any* k. The proof for $k \geq 2$ is similar, except that now it is necessary to require that ϕ be on shell (i.e., ϕ must satisfy the Euler–Lagrange equations) for (II.6) to hold. ♦

Remark II.4. Although this theorem presupposes that one has chosen an embedding $\mathrm{Diff}_c(X) \to \mathcal{G}$, it is shown in Gotay and Marsden [1992] that this choice is ultimately irrelevant. Thus even though the last two terms in (II.10) individually depend upon the choice of embedding, their sum does not. See the recent review paper by Forger and Römer [2004] for an alternative approach which avoids this issue. ♦

Remark II.5. As noted previously, the image of $\mathcal{G}$ under the natural projection $\mathrm{Aut}(Y) \to \mathrm{Diff}(X)$ may be strictly larger than $\mathrm{Diff}_c(X)$. Let us denote this image by $\mathcal{D}\,(= \mathcal{G}/\mathcal{G}_{\mathrm{Id}})$. In such cases one could consider an embedding $\mathcal{D} \to \mathcal{G}$. But then the relation (II.6) between the integrals of $J^{\mathcal{L}}$ and $\mathfrak{T}^\mu{}_\nu$ for general vector fields ξ in the Lie algebra of $\mathcal{D}$ might only hold modulo surface terms. For instance, suppose that $k = 1$ and that $\mathcal{D} = \mathrm{Diff}(X)$. Tracing back through the proof of Theorem II.1, and using Stokes' theorem in place of the divergence theorem, we obtain

$$\int_\Sigma i_\Sigma^*(j^1\phi)^* J^{\mathcal{L}}(\xi_Y) = \int_\Sigma \mathfrak{T}^\mu{}_\nu \xi^\nu d^n x_\mu + \frac{1}{2}\int_{\partial\Sigma} L_A{}^\mu C^{A\rho}{}_\nu \xi^\nu d^{n-1}x_{\mu\rho}$$

for $\xi \in \mathfrak{X}(X)$. In the asymptotically flat context, it is plausible that surface terms such as the one in this expression could be identified with the energy and momentum of the gravitational field. (See the gravity example in Part V for further discussion of this point.)

In any event we emphasize that Theorem II.1 *as stated* is valid regardless of what $\mathcal{D}$ is, as long as $\mathrm{Diff}_c(X) \subset \mathcal{D}$, and thus can always be used to *define* $\mathfrak{T}^\mu{}_\nu$. ♦

We now collect some of the important properties of our SEM tensor density. Proofs not given below may be found in Gotay and Marsden [1992].

First, combining (II.10) and the last condition in (II.5) we obtain the ***generalized Hilbert formula***

$$\mathfrak{T}^{\mu}{}_{\nu} = -C^{A\mu}{}_{\nu}\frac{\delta L}{\delta y^A}. \tag{II.11}$$

In particular, we conclude that $\mathfrak{T}(\phi)$ vanishes on shell. This is a typical feature of parametrized theories in which *all* fields are variational; compare Remark 6E.18. However, this is not the full story; see the discussion after Proposition II.8 as to what happens when only the "matter fields" are variational.

Proposition II.6.

$$D_\mu \mathfrak{T}^{\mu}{}_{\nu} = (v^A{}_\nu - C^A{}_\nu)\frac{\delta L}{\delta y^A}.$$

See Proposition II.10 following for an application of this result.

Proof. From (II.9)

$$\begin{aligned} D_\mu \mathfrak{T}^{\mu}{}_{\nu} &= D_\mu\left(L\delta^\mu{}_\nu - L_A{}^\mu v^A{}_\nu + K^\mu_\nu + D_\rho K^{\mu\rho}_\nu\right) \\ &= \big(L_{,\nu} + L_A v^A{}_\nu + L_A{}^\mu v^A{}_{\mu\nu} - (D_\mu L_A{}^\mu)v^A{}_\nu - L_A{}^\mu v^A{}_{\nu\mu} \\ &\quad + D_\mu K^\mu_\nu + D_\mu D_\rho K^{\mu\rho}_\nu\big)\,. \end{aligned}$$

Here the third and fifth terms cancel while the second and fourth combine to produce a variational derivative. Using the first of (II.5) the last term is seen to vanish. Thus we obtain

$$D_\mu \mathfrak{T}^{\mu}{}_{\nu} = L_{,\nu} + D_\mu K^\mu_\nu + \frac{\delta L}{\delta y^A}v^A{}_\nu.$$

The result now follows from the last of (II.5). ∎

Proposition II.7. $\mathfrak{T}(\phi)$ *is gauge-covariant.*

By this we mean that the tensor density $\mathfrak{T}(\phi)$ satisfies

$$\mathfrak{T}(\eta_{\mathcal{Y}}(\phi)) = \eta_{X*}\mathfrak{T}(\phi) \tag{II.12}$$

for all $\eta \in \mathcal{G}$ and solutions ϕ of the Euler–Lagrange equations. When the gauge transformation η is purely "internal", i.e., $\eta \in \mathcal{G}_{\mathrm{Id}}$, (II.12) reduces to $\mathfrak{T}(\eta_{\mathcal{Y}}(\phi)) = \mathfrak{T}(\phi)$. Thus $\mathfrak{T}(\phi)$ is actually gauge-*in*variant in this case. At the opposite extreme, when η_Y is the lift of a diffeomorphism of X, (II.12) reiterates the fact that $\mathfrak{T}(\phi)$ is a tensor density.

One can also inquire as to the dependence of $\mathfrak{T}^{\mu}{}_{\nu}$ upon the choice of Lagrangian density $\mathcal{L}$. As is well known, $\mathcal{L}$ is not uniquely fixed; one is free to add a $\mathcal{G}$-equivariant divergence to it without changing the physical content of the theory. But $\mathfrak{T}^{\mu}{}_{\nu}$, unlike $\mathfrak{t}^{\mu}{}_{\nu}$, is independent of such ambiguities, as the next result shows.

Proposition II.8. *$\mathfrak{T}^{\mu}{}_{\nu}$ depends only upon the divergence equivalence class of $\mathcal{L}$.*

We now turn to a brief discussion of what happens when a metric g is present on X. This metric may or may not be variational and, in any case, need not represent gravity. We refer to the remaining fields ψ^a collectively as "matter" fields (even though this appellation may be somewhat inappropriate in particular examples). Below $\mathcal{L}$ will always refer to the matter Lagrangian; a "free field" Lagrangian for the metric is immaterial. We assume that the ψ^a are on shell.

Let $\mathrm{Diff}_c(X)$ act on g by pushforward. Then for $\xi \in \mathfrak{X}_c(X)$ the "metric component" of ξ_Y is

$$(\xi_Y)_{\alpha\beta} = -(g_{\nu\beta}\xi^{\nu}{}_{,\alpha} + g_{\nu\alpha}\xi^{\nu}{}_{,\beta});$$

hence

$$C_{\alpha\beta}{}^{\rho}{}_{\nu} = -(g_{\nu\beta}\delta^{\rho}{}_{\alpha} + g_{\nu\alpha}\delta^{\rho}{}_{\beta}) \tag{II.13}$$

is the only nonzero coefficient in (II.4).

Formula (II.10) gives

$$\begin{aligned}\mathfrak{T}^{\mu}{}_{\nu} &= L\delta^{\mu}{}_{\nu} - L_a{}^{\mu}v^a{}_{\nu} + L_a{}^{\mu}C^a{}_{\nu} + D_\rho(L_a{}^{\rho}C^{a\mu}{}_{\nu}) \\ &\quad - \frac{\partial L}{\partial w_{\alpha\beta\mu}}w_{\alpha\beta\nu} - 2D_\rho\left(\frac{\partial L}{\partial w_{\mu\beta\rho}}g_{\nu\beta}\right)\end{aligned} \tag{II.14}$$

where $w_{\alpha\beta\nu}$ are the velocity variables associated to the $g_{\alpha\beta}$. On the other hand, from (II.11) and (II.13) we have

$$\mathfrak{T}^{\mu}{}_{\nu} = -C_{\alpha\beta}{}^{\mu}{}_{\nu}\frac{\delta L}{\delta g_{\alpha\beta}} = 2\frac{\delta L}{\delta g_{\mu\rho}}g_{\nu\rho}.$$

Raising indices then yields

Theorem II.9.

$$\mathfrak{T}^{\mu\nu} = 2\frac{\delta L}{\delta g_{\mu\nu}}. \tag{II.15}$$

As a consequence, the "matter" SEM tensor density $\mathfrak{T}^{\mu\nu}$ is manifestly symmetric, gauge-covariant, and independent of the choice of embedding—attributes that are hardly obvious from (II.14)! Nor does it vanish in general. Perhaps most importantly, this result imparts a straightforward physical interpretation à la Noether to the Hilbert SEM tensor.

We derive one last consequence of our formalism.

Proposition II.10. *$\mathfrak{T}^{\mu}{}_{\nu}$ is covariantly conserved.*

Proof. Fix a point $x \in X$ and work in normal coordinates centered there. Since g has pure index 1, Proposition II.6 gives

$$\nabla_{\mu}\mathfrak{T}^{\mu}{}_{\nu}(x) = D_{\mu}T^{\mu}{}_{\nu}(x) = \left(\frac{\delta L}{\delta g_{\alpha\beta}} w_{\alpha\beta\nu}\right)(x).$$

But $w_{\alpha\beta\nu}(x) = 0$ in normal coordinates centered at x. ■

The proof of this result does not rely on the field equations for the metric. (Indeed, g may not even *have* field equations.) In particular, it is not necessary to couple a field theory to gravity and appeal to Einstein's equations to force $\mathfrak{T}^{\mu}{}_{\nu}$ to be covariantly conserved; Proposition II.10 is a general property of $\mathfrak{T}^{\mu}{}_{\nu}$. See Fischer [1982,1985] for further discussion of this and related matters.

Examples

a Particle Mechanics. For parametrized particle mechanics, we have $J^{\mathfrak{L}}(\chi) = -E\chi$ for $\chi \in \mathcal{F}(\mathbb{R})$. Thus the (scalar) SEM tensor reduces to just $\mathfrak{T}^{t}{}_{t} = -E$ and vanishes on shell, consistent with Example 4D.**a**.

b Electromagnetism. The Maxwell theory provides a simple, yet nontrivial example of our constructions.

Since $\partial L/\partial v_{\alpha\beta} = F^{\alpha\beta}\sqrt{-g}$, the canonical SEM tensor (II.1) is

$$\mathfrak{t}^{\mu\nu} = -\left[\frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + g^{\nu\beta}F^{\alpha\mu}D_{\beta}A_{\alpha}\right]\sqrt{-g}.$$

It is clearly neither symmetric nor gauge-invariant. To "fix" it, one adds the term $g^{\nu\beta}F^{\alpha\mu}D_{\alpha}A_{\beta}\sqrt{-g}$, thereby producing

$$\mathfrak{T}^{\mu\nu} = -\left[\frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + g^{\nu\beta}F^{\alpha\mu}F_{\beta\alpha}\right]\sqrt{-g}. \tag{II.16}$$

We now derive (II.16) using our methods. The standard embedding of $\mathrm{Diff}_c(X)$ into the gauge group $\mathcal{G} = \mathrm{Diff}(X) \ltimes C^\infty(X)$ is simply $\eta \mapsto (\eta, 0)$. However, to illustrate the independence of our results upon the choice of embedding, we instead choose the nonstandard one given as follows: fix $\chi \in \mathcal{F}(X)$ and define $\eta \mapsto (\eta, \eta_*\chi - \chi)$. Then the induced action of $\mathrm{Diff}_c(X)$ on Y is

$$\eta_Y \cdot A = \eta_* A + \mathbf{d}(\eta_*\chi - \chi).$$

Then for $\xi \in \mathfrak{X}_c(X)$,

$$\xi_Y = \xi^\nu \frac{\partial}{\partial x^\nu} - (A_\nu \xi^\nu{}_{,\alpha} + \chi_{,\nu}\xi^\nu{}_{,\alpha} + \chi_{,\alpha\nu}\xi^\nu)\frac{\partial}{\partial A_\alpha}$$

so (II.4) holds with

$$C_{\alpha\nu}^{\ \ \rho} = -(A_\nu + \chi_{,\nu})\delta^\rho{}_\alpha \quad \text{and} \quad C_{\alpha\nu} = -\chi_{,\alpha\nu}.$$

Thus $k = 1$, but note that zeroth order derivative terms appear as well.

Applying (II.10), and observing that the metric is nonderivatively coupled to A, we obtain

$$\begin{aligned}\mathfrak{T}^\mu{}_\nu = &- \left[\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\delta^\mu{}_\nu + F^{\alpha\mu}D_\nu A_\alpha\right]\sqrt{-g} \\ &- F^{\alpha\mu}\chi_{,\alpha\nu}\sqrt{-g} - D_\alpha\left(F^{\mu\alpha}[A_\nu + \chi_{,\nu}]\sqrt{-g}\right).\end{aligned}$$

Using Maxwell's equations $D_\alpha(F^{\mu\alpha}\sqrt{-g}) = 0$, the last term becomes

$$-F^{\mu\alpha}(D_\alpha A_\nu + \chi_{,\nu\alpha})\sqrt{-g}.$$

Substituting into the above and raising the index we obtain (II.16). All trace of χ has disappeared, as it must by the general theory presented.

Note that (II.16) does not necessarily vanish when Maxwell's equations are satisfied, because the metric is nonvariational—see Example 4D.**b**.

c A Topological Field Theory. We compute

$$\mathfrak{T}^\mu{}_\nu = \epsilon^{\mu\alpha\beta}A_\nu F_{\alpha\beta}.$$

That $\mathfrak{T}^\mu{}_\nu$ vanishes on shell reflects the fact that topological field theories have no "local physics," and hence no localizable "energy" or "momentum." (It is necessary to qualify these terms since, in the absence of a metric, one cannot distinguish one from the other.) However, this does not preclude the possibility that such a theory has a total nonzero energy/momentum content. In fact, using the formula in Remark II.5, we can explicitly compute the "topological" energy/momentum at infinity:

$$-\int_\Sigma i_\Sigma^*(j^1A)^* J^{\mathcal{L}}(\xi_Y) = \int_{\partial\Sigma} [\mathbf{i}(\xi)A]A$$

for arbitrary $\xi \in \mathfrak{X}(X)$.

d Bosonic Strings. Since h is nonderivatively coupled to ϕ, (II.10) gives

$$\mathfrak{T}_{\mu\nu} = \left[g_{AB} v^A{}_\mu v^B{}_\nu - \frac{1}{2} h_{\mu\nu} h^{\alpha\beta} g_{AB} v^A{}_\alpha v^B{}_\beta \right] \sqrt{-h}.$$

As ϕ is a scalar field, this coincides with the canonical SEM tensor density. Again we see that the SEM tensor density vanishes on shell. One can also check directly that this formula is consistent with Theorem II.9:

$$\mathfrak{T}_{\mu\nu} = -2 \frac{\delta L}{\delta h^{\mu\nu}}.$$

In the jargon of elasticity theory, our treatment of the string corresponds to the "body" representation. This is consistent with $\mathfrak{T}_{\mu\nu}$ being a 2×2 matrix; in some sense it describes the "internal" distribution of energy, momentum and stress. Physically, then, the vanishing of $\mathfrak{T}_{\mu\nu}$ on shell can be interpreted to mean that if the string is to be harmonically mapped into spacetime, then it must be "internally unstressed." To obtain the "physical" 4×4 SEM tensor on the spacetime (M, g), one would have to work in the inverse material representation and consider a "cloud" of strings.

In this example, the parameter space X was not the physical spacetime. This happens in other contexts as well (including our Example **a**) . For instance, Künzle and Duval [1986] have constructed a Kaluza-Klein version of classical field theory; in their approach X is a certain circle bundle over spacetime. The SEM tensor is then defined on this space and so is a 5×5 matrix; in the case of an adiabatic fluid, the additional components of $\mathfrak{T}^\mu{}_\nu$ can be interpreted as an entropy-flux vector (Künzle [1986]).

Remark II.11. String theory has a certain communality with relativistic elasticity. (See Beig and Schmidt [2003] for an introduction to this theory.) Relativistic elasticity has, as its basic fields, the world tube of the elastic material and a Lorentz metric g on the physical spacetime M. The world tube is viewed as a map $\Phi : \mathcal{B} \times \mathbb{R} \to M$, where $\mathcal{B}$ is a 3-dimensional reference region and $\mathbb{R}$ is the time axis. The matter Lagrangian density is a given constitutive function of the relativistic spatial version of the Cauchy-Green tensor, namely $g + u \otimes u$, where u is the world velocity field of Φ. See Marsden and Hughes [1983], §5.7, for more details on how this is set up. The spacetime diffeomorphism group acts on the world tube by composition on the left, which is a relativistic version of the principle of material frame indifference, and it acts on the metric as usual, by push-forward. Thus, our theory applies to this case, and so one must have the stress energy momentum tensor given by either via the Noether based definition (II.6), or equivalently via the Hilbert formula (II.15). The former seems not to be known in relativistic elasticity. The Hilbert formula is common in the literature and is found on page 313 of Marsden and Hughes [1983], which can also be consulted for additional references. ♦

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 Space-time.* 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.