JGP - Vol. 4, n. 2, 1987

Variational principles on principal fiber bundles: A geometric theory of Clebsch potentials and Lin constraints

HERNAN CENDRA

Universidad Nacional del sur - Dep. de Matemática
Ave. Alem 1253, 8000 Bahía Blanca, Argentina

ALBERTO IBORT

Departmento de Fisica Téorica, Universidad de Zaragoza
50009 Zaragoza, Spain

JERROLD MARSDEN

Department of Mathematics, University of California, Berkeley
Berkeley, California, 94720, U.S.A.

Abstract. *The geometric theory of Lin constraints and variational principles in terms of Clebsch variables proposed recently by Cendra and Marsden [1987] will be generalized to include those systems defined not only on configuration spaces which are products of Lie groups and vector spaces but on configuration spaces which are principal bundles with structural group G. This generalization includes, for example, fluids with free boundaries, Yang-Mills fields, and it will be very useful, as it will be shown later, to illustrate some aspects of the theory of particles moving in a Yang-Mills field in both its variational and Hamiltonian aspects.*

1. INTRODUCTION

The origin and necessity of Lin constraints can be easily understood by considering a Lagrangian L defined on the tangent bundle of a trivial principal bundle $B \times G$, and assuming that L is invariant under the action of the group G lifted to the tangent bundle $T(B \times G)$. Thus L defines a Lagrangian $L' : TB \times \mathfrak{g} \to \mathbf{R}$. One

Key-Words: Variational Principles, Principal Bundles, Yang-Mills fields.
1980 MSC: 53C, 58 E 30, 70 H.

Research partially supported by NSF Contract DMS 87 - 02502.

discovers that it is impossible to get the equations of motion by using a naive variational principle associated to L', for example, varying curves $(x(t), v(t))$ on $B \times \mathfrak{g}$ with the usual fixed endpoint condition. For instance, for the rigid body with $G = SO(3)$ and B a single point, L' is the usual kinetic energy in body representation and the naive variational principle for L' does *not* give Euler's equations. The first method proposed to overcome this difficulty was the «Clebsch representation» technique, which adds more variables, extending the space of variables (x, v) and introduces a new Lagrangian defined on the extended space, which will give the correct equations of motion. One such extension is to $T(B \times G)$ itself, but there are others as well. It is plausible that if we use the same idea in the context of configuration spaces which are nontrivial principal bundles we should describe separately the variations along the vertical («group») directions and the horizontal («base») directions, or equivalently we would have to choose a connection to separate degrees of freedom along the fiber directions from the degrees of freedom along the base directions. Obviously such a choice may have global consequences in the description of the reduced dynamics in the same way that the equations of motion for a particle moving in a Yang-Mills field depends on the connection we choose in the gauge fiber bundle of the system (see Montgomery [1984] and references therein). This example using these ideas will be discussed in §5. In §2 we describe some notations and ideas about principal and associated fiber bundles. In §3 we introduce the horizontal Lin constraints and a related variational principle. In §4 we use this variational principle to find equations of motion for invariant Lagrangians.

2. ASSOCIATED BUNDLES AND CLEBSCH SPACES

Let $\pi : P \to B$ be a principal bundle with structure group G acting on the right. Suppose $\rho : G \times M \to M$ is a (left) action of G on a manifold M. Recall that the *associated bundle* with fiber M is $P \times_G M = (P \times M)/\sim$, where the equivalence relation $\sim$ is defined by $(p_1, m_1) \sim (p_2, m_2)$ if and only if $p_1 = p_2 g$ and $m_1 = \rho(g^{-1}, m_2)$, for some $g \in G$. The equivalence class $[(p, m)]$ will be often written pm. We have the following commutative diagram, where the meaning of the arrows is the obvious one:

$$\begin{array}{ccc} P \times M & \xrightarrow{\pi_1} & P \\ {\scriptstyle proj}\downarrow & & \downarrow{\scriptstyle \pi} \\ P \times_G M & \xrightarrow[\pi_M]{} & B \end{array}$$

For each $p \in P$, the map $i_p : M \to P \times_G M$ defined by $i_p(m) = pm$ is an embedding. For simplicity, we will write: $(pm, p\dot{m}) := p\dot{m} := T_m i_p(m, \dot{m})$. Likewise for each $m \in M$, we have the map $i_m : P \to P \times_G M$, $i_m(p) = pm$. Also $Ti_m(p, \dot{p}) := (pm, \dot{p}m) := \dot{p}m$ is a convenient notation. Thus, if $(p(t), m(t))$ is a curve in $P \times M$, the tangent vector to the curve $p(t)m(t)$ is given by: $T_p i_m(\dot{p}) + T_m i_p(\dot{m}) = \dot{p}m + p\dot{m}$. Given a connection A on P, a parallel transport operation is induced on $P \times_G M$. The horizontal subspace of $T_{pm}(P \times_G M)$ is defined to be the subspace $H^M{}_{pm}$ spanned by tangent vectors to curves of the form $p(t)m$ with $p(t)$ horizontal in P (i.e., $A(\dot{p}) = 0$) and $m \in M$ fixed. The horizontal distribution H^M is the kernel of the $T(P \times_G M)$-valued 1-form A^M defined as follows:

$$A^M{}_{pm}(X_{pm}) = X_{pm} - X^h{}_{pm} = X^v{}_{pm}$$

where $X_{pm} \in T_{pm}(P \times_G M)$, $X^h{}_{pm}$ is the horizontal projection of X_{pm}, and $X^v{}_{pm}$ is the vertical projection of X_{pm}; (i.e. the projection of the tangent space to the fiber $\pi_M{}^{-1}(\pi, (p)) \subset P \times_G M$ or, in other words, $X^v{}_{pm} \in V^M{}_{pm} = T_{pm}\pi_M{}^{-1}(\pi(p))$. It is easy to see that for fixed p, the curve $pm(t)$ is vertical and also that the horizontal component of

$$\left.\frac{dp(t)m}{dt}\right|_{t=t_0} = X_{pm}$$

is $\dot{p}^h(t_0)m$ where $p^h(t)$ is the horizontal lift of $p(t)$ such that $p^h(t_0) = p(t_0)$. More generally, for a given curve $(p(t), m(t)) \in p \times M$, the horizontal component of the tangent vector X_{pm} is $\dot{p}^h(t_0)\, m(t_0)$. Thus the vertical component of

$$\frac{dp(t)m(t)}{dt} = X$$

is

$$\begin{aligned}\dot{p}m + p\dot{m} - \dot{p}^h m &= (\dot{p} - \dot{p}^h)m + p\dot{m} = \dot{p}^v m + p\dot{m}\\ &= (pA(\dot{p}))m + p\dot{m} = p(A(\dot{p})m) + p\dot{m}\end{aligned}$$

where $pA(\dot{p})$ is the infinitesimal generator of $A(\dot{p}) \in \mathfrak{g}$ calculated at $p \in P$, and the last equality comes from the formula: $(pu)m = p(um)$ for $u \in \mathfrak{g}$, which in turn is the infinitesimal version of the equality $(pg)m = p(gm)$. Thus we can write

$$A^M\left(\frac{dpm}{dt}\right) = p\dot{m} + (pA(\dot{p}))m = p\dot{m} + p(A(\dot{p})m)$$

There is a canonical inclusion $P \times_G TM \subset T(P \times_G M)$. Notice also that A^M is actually $P \times_G TM$ - valued; in fact both pm and $p(A(p)m)$ belong to $P \times_G TM$.

3. HORIZONTAL LIN CONSTRAINTS AND VARIATIONAL PRINCIPLES

As in the previous section, let $\pi : P \to B$ be a principal bundle with structure group G, and assume that $\rho : G \times M \to M$ is a given left action. In addition, assume that the following condition is satisfied from now on:

G *There is a G-invariant open set $U \subset M$ such that G is embedded into its orbit Gm for each $m \in U$.*

The assumption that U is open should be properly interpreted in each example, especially in infinite dimensional cases. This involves issues of functional analysis that we will not detail here. The point to keep in mind is that U should allow enough variations of curves to apply the usual variational techniques. Similarly, the assertion that G is embedded onto its orbit should be properly interpreted in examples. In a number of examples, we can choose U to be G itself. To handle the notations about variational principles on nontrivial principal bundles that appear in the statement of the main result of this section, we introduce some notation.

We will be dealing with several spaces of curves. A curve in P will be generally denoted $p:[t_0, t_1] \to P$, while fixed points on P will be denoted $p_0, p_1, p_2, \ldots$ etc. Thus, the condition $p(t_0) = p_0$, means that the left endpoint of the curve p is p_0, and the condition $p(t_1) = p_1$ means that the right endpoint of the curve p is p_1. The condition for curves of having fixed endpoints are of common use in variational techniques; for instance, in Hamilton's Variational Principle, variations δp are allowed such that $\delta p(t_0) = 0$, $\delta p(t_1) = 0$. More precisely, if $p(t)$ is a given curve in P satisfying $p(t_i) = p_i$, $i = 1, 2$ a *variation* of the curve p is a C^∞ family of curves $p(t, \lambda)$ such that $p(t, 0) = p(t)$, and the condition $\delta p(t_i) = 0, i = 0, 1$, corresponds to $p(t_i, \lambda) = p_i$, for $i = 0, 2$ and all λ. The following notation will be useful in the rest of this paper. For $p_0, p_1 \in P$, define

$$\Omega(P; p_0) = \{p : [t_0, t_1] \to P \mid p(t_0) = p_0\}$$

$$\Omega(P; p_0, p_1) = \{p : [t_0, t_1] \to P \mid p(t_0) = p_0, p(t_1) = p_1\}.$$

Likewise, for fixed $m_0, m_1 \in M$, define

$$\Omega(M; m_0) = \{m : [t_0, t_1] \to M \mid m(t_0) = m_0\}.$$

and

$$\Omega(M; m_0, m_1) = \{m : [t_0, t_1] \to M \mid m(t_i) = m_i,\ i = 0, 1\}.$$

For $p_0 \in P$ and $x_1 \in B$, define

$$\Omega(P; p_0, x_1) = \{ p : [t_0, t_1] \to P \mid p(t_0) = p_0,\ \pi(p(t_0)) = x_1 \}.$$

In other words, $\Omega(P; p_0, x_1)$ is the manifold of all curves p having fixed left endpoint p_0 and also having the right endpoint «vertically free», which means that only the projection of the right endpoint is fixed to be $x_1 \in B$. For instance if $P = B \times G$ is a trivial bundle and $p_0 = (x_0, g_0) \in P$ and $x_1 \in B$ are fixed, curves $p(t)$ belonging to $\Omega(P; p_0, x_1)$ are of the type $p(t) = (x(t), g(t))$, where $g(t_0) = g_0$ and $x(t_i) = x_i, i = 0, 1$.

Fix a connection A on the principal bundle P. Then we have the notion of horizontality for curves p, with respect to the connection A. Given any curve $x(t)$ on B, such that $x(t_0) = x_0$, there is a unique «horizontal lift» of $x(t)$, say $p(t)$, such that $p(t_0) = p_0$, and $\pi(p(t)) = x(t)$ for all $t \in [t_0, t_1]$. Observe that if $x(t_1) = x_1$, then the curve $p(t)$ belongs to $\Omega(P; p_0, x_1)$. Thus, given any curve $p \in \Omega(P; p_0)$ there is a unique decomposition $p(t) = p^h(t) \cdot g^p(t)$, where $p^h(t)$ is a horizontal curve on P satisfying $p^h(t_0) = p_0$ and $g^p(t)$ is a curve on G satisfying $g^p(t_0) = e$. In fact, $p^h(t)$ is the horizontal lifting of $\pi(p(t)) = x(t)$, having left endpoint p_0. Another useful notation is the following. Let $p \in \Omega(P; p_0, x_1)$ and $m_0 \in M$ be given. Then there is a unique curve $m^p(t)$ belonging to $\Omega(M; m_0)$ such that $p(t)m^p(t)$ is horizontal and $p(t_0)m^p(t_0) = p_0 m_0$. This curve is defined by

$$m^p(t) = [g^p(t)]^{-1} m_0.$$

In fact, we have

$$p(t)m^p(t) = p^h(t)[g^p(t)]^{-1} m_0 = p^h(t) m_0$$

which proves that $p(t)\, m^p(t)$ is horizontal, as a curve on $P \times_G M$, and on the other hand it is also easy to see that $p(t_0)\, m^p(t_0) = p_0 m_0$. Assumption **G** implies that if $m_0 \in U$ and m^p is defined as before, then $m^p(t) \in U$ for all $t \in [t_0, t_1]$.

With $p_0, p_1 \in P$, $\pi(p_1) = x_1 \in B$, and $m_0 \in M$ fixed, define the functions

$$m_1 : \Omega(P; p_0, x_1) \to U \quad \text{by} \quad m_1(p) = i_{p_1}^{-1}(p^h(t_1) m_0)$$

and

$$g_1 : \Omega(P; p_0, x_1) \to G \quad \text{by} \quad p^h(t_1) = p_1 g_1(p).$$

These definitions make sense since both $p^h(t_1)$ and p_1 lie over the same base point x_1. Notice especially that while these two maps are defined on the space of curves with right hand endpoint over the base point x_1, the maps themselves depend on the choice of p_1. To get an intuitive picture of the previous notions, let us explicitly describe them for the case of a trivial bundle $P = B \times G$ with A the trivial connection; i.e. horizontal curves on P have the form $p(t) = (x(t), g_0)$ where $g_0 \in G$ is independent of t. Given $p_0 = (x_0, g_0)$, $p_1 = (x_1, g_1)$, $m_0 \in M$ and a curve $p \in \Omega(P; p_0, x_1)$, say $p(t) = (x(t), g(t))$, we can give formulas for

$p^h, g^p, m^p, m_1 : \Omega(P; p_0, x_1) \to U$ and $g_1 : \Omega(P; p_0, x_1) \to G$ as follows. First observe that we have $P \times_G M \cong B \times M$ via $[(x, g), m] \mapsto (x, gm)$. This implies that horizontal curves on $P \times_G M$ are of the form $(x(t), m_0)$, where $m_0 \in M$ is independent of t. Then we have

$$p^h(t) = (x(t), g_0), g^p(t) = [g_0]^{-1} g(t),$$
$$m^p(t) = [g(t)]^{-1} g_0 m_0 \quad \text{and} \quad m_1(p) = [g_1]^{-1} g_0 m_0$$

where $g_1 \in G$ is the second component of $p_1 = (x_1, g_1)$ which was given. (There is a slight abuse of notation, since g_1 is also used to denote a function $\Omega(P; p_0, x_1) \to G$). This can be checked as follows. From the definition we have: $p_1 m_1(p) = p^h(t_1) m_0$. Since $p_1 m_1(p) = (x_1, g_1 m_1(p))$ and $p^h(t_1) m_0 = (x_1, g_0 m_0)$ we conclude that $m_1(p) = [g_1]^{-1} g_0 m_0$. For the particular case of a trivial connection $m_1(p)$ only depends on the given values of p_0 and p_1 and it is really independent of the curve p. The meaning of this will become clearer later on. This is *not* the case for a nontrivial connection. Finally we have

$$g_1(p) = [g_1]^{-1} g_0$$

which shows that for a trivial connection, $g_1(p)$ is a constant i.e. does not depend on p and only depends on the fixed values of p_0, p_1.

In the case of a general bundle P and a general connection A, we can check as an easy consequence of the definitions that

$$m_1(p) = g_1(p) m_0.$$

The purpose of the previous definitions is the following lemma.

LEMMA 3.1. *Let $p_0, p_1 \in P$ and $m_0 \in U$ be given:*
(a) *If $p \in \Omega(P; p_0, x_1)$, $m \in \Omega(U, m_0, m_1(p))$ and $p(t)m(t)$ is horizontal, then $p(t_1) = p_1$. (Recall that m_1 depends on the choice of the point p_1).*
(b) *Conversely, given $p \in \Omega(P; p_0, p_1)$, there is a unique $m \in \Omega(U, m_0, m_1(p))$ such that pm is horizontal; in fact $m = m^p$ satisfies this requirement.*

Before giving the proof, we comment on the meaning of this lemma in the particular case of a trivial bundle with a trivial connection. Given $p(t) = (x(t), g(t))$, and $m(t)$, the condition of horizontality for $p(t)m(t) = (x(t), g(t)m(t))$ means simply that $g(t)m(t) = g(t_0)$ is independent of t. Thus if $p(t_0) = (x_0, g_0)$, $m(t_0) = m_0$ and $p(t)m(t)$ is horizontal then $g(t_1)m(t_1) = g_0 m_0$. Since we are also assuming that $m(t_1) = m_1(p) = [g_1]^{-1} g_0 m_0$ according to the previous discussion, we get $g(t_1)\,[g_1]^{-1} g_0 m_0 = g_0 m_0$, and since $m_0 \in U$, condition **G** together with the previous equality implies $g(t_1) = g_1$, and since, by assumption, $x(t_1) = x_1$, we get $p(t_1) = (x(t_1), g(t_1)) = (x_1, g_1) = p_1$, which proves part (a) of the lemma in the particular case of a trivial connection. Part (b) can be proved in a similar way.

The essential content of the lemma, in the case of a trivial connection, consists of the assertion that $g(t)m(t) =$ constant (horizontality) and $m(t_1) = \text{const} \in U$ (i.e. $m(t_1) = m_1(p)$) together imply that $g(t_1)$ is also a constant. In the nontrivial case however, difficulties arise precisely because $m_1(p)$ is not constant (independent of p), and could not possibility be chosen to be a constant, and it is interesting that nevertheless, the lemma, as stated is still valid.

Proof of Lemma 3.1. Let us prove the following equivalences, which are valid for curves $m \in \Omega(U, m_0)$, $p \in \Omega(P; p_0, x_1)$ as before and p_0, p_1, m_0 are fixed.

$$p_1 = p(t_1) \Longleftrightarrow [g_1(p)]^{-1} = g^p(t_1) \Longleftrightarrow m^p(t_1) = m_1(p).$$

In fact, we have

$$p_1 = p(t_1) \Longleftrightarrow p^h(t_1)[g_1(p)]^{-1} = p^h(t_1)g^p(t_1) \Longleftrightarrow [g_1(p)]^{-1} = g^p(t_1)$$

where the last equivalence comes from the fact that the action of G on P is free. On the other hand, using formulas before the lemma and Assumption **G** we conclude that

$$m^p(t_1) = m_1(p) \Longleftrightarrow [g_1(p)]^{-1} = g^p(t_1).$$

Using these facts we can prove the lemma as follows

$$pm \text{ horizontal and } m(t_1) = m_1(p) \Longleftrightarrow m = m^p \text{ and } m(t_1) = m_1(p)$$

$$\Rightarrow m^p(t_1) = m_1(p) \Rightarrow p_1 = p(t_1).$$

Conversely, let $p \in \Omega(P; p_0, p_1)$ be given. Then $p(t_1) = p_1$ and therefore $m^p(t_1) = m_1(p)$. Thus if we choose $m = m^p$, then p^m is horizontal and $m \in \Omega(U; m_0, m_1(p))$. ∎

We can write the previous lemma in a more compact form as follows. Recalling that m_1 depends implicitly on the choice of p_1, and we let $x_1 = \pi(p_1)$, define the sets

$$\Omega(P \times M; p_0, p_1, m_0) = \{(p, m) : [t_0, t_1] \to P \times M \mid p(t_0) = p_0,$$
$$\pi(p(t_1)) = x_1, m(t_0) = m_0, m(t_1) = m_1(p)\}$$

and $\Omega^H(P \times M; p_0, p_1, m_0)$ the subset of horizontal curves, i.e. curves satisfying

$$A^M\left(\frac{\mathrm{d}pm}{\mathrm{d}t}\right) = 0.$$

Then we have the following rewording of lemma 3.1.

COROLLARY 3.2. *The map*

$$\Omega^H(P \times M; p_0, p_1, m_0) \to \Omega(P; p_0, p_1)$$

defined by $(p, m) \to p$, *is an isomorphism onto. The inverse is given by* $p \to (p, m^p)$.

The main purpose of this section is to introduce the horizontal constraint given by

$$A^M\left(\frac{\mathrm{d}pm}{\mathrm{d}t}\right) = 0$$

via a Lagrange multiplier to get a variational principle involving Clebsch potentials and the variable $v = A(\dot{p})$ as quantities to be varied independently. We represent by (m, α) or, sometimes, simply $\alpha = \alpha_m$ an element of $T^*_m M$. Thus elements of $P \times_G T^*M$ will be denoted simply $p\alpha_m = p\alpha$ (the action of G on T^*M is the cotangent lifting of ρ). We have a natural pairing:

$$\langle , \rangle : P \times_G TM \oplus P \times_G T^*M \to \mathbf{R}$$

where $\oplus$ stands for the Whitney sum over the basis $P \times_G M$. We define $\langle , \rangle$ by the formula: $\langle p\dot{m}, p\alpha \rangle = \langle \dot{m}, \alpha \rangle$, and we can check that this is well defined. It is sometimes useful to work with a canonical trivialization of T^*M say $M^1 \times F^*$, where $M^1 \subset M$ is open and F is isomorphic to the tangent space to M at some point belonging to M^1. Thus for given (m, α) belonging to this trivialization, we can interpret m as being an element of M^1 and α as being an element of F^*. We still need some more notation to be used in the following theorem. As before, assume $p_0, p_1 \in P$ are fixed, $x_1 = \pi(p_1)$ and $m_0 \in M$ is fixed. Define

$$\Omega(P \times T^*U; p_0, p_1, m_0) = \{(p; m, \alpha) : [t_0, t_1] \to P \times T^*U \mid p(t_0) = p_0, \\ \pi(p(t_1)) = x_1, m(t_0) = m_0, m(t_1) = m_1(p)\}.$$

Notice that for each given curve p on P satisfying $p(t_0) = p_0$ and $\pi(p(t_1)) = x_1$, the point $m_1(p)$ remains fixed. Now choose any curve m on M satisfying $m(t_0) = m_0$ and $m(t_1) = m_1(p)$. Finally, choose the curve (m, α) on T^*M arbitrarily, except that the component m satisfies the previous endpoint conditions. Then $(p; m, \alpha)$ is a typical element of $\Omega(P \times T^*U; p_0, p_1, m_0)$. Notice that the point

p_1 is implicit in the definition of m_1. The horizontal constraint on curves (p, m), i.e., the condition that pm be horizontal, is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left\langle A^M\left(\frac{\mathrm{d}pm}{\mathrm{d}t}\right), \alpha_\lambda\right\rangle\Big|_{\lambda=0} = 0$$

for all variations α_λ of the curve α. Arbitrarily variations (p, m, α_λ) of the curve (p, m, α) are allowed in the manifold of curves $\Omega(P \times T^*U; p_0, p_1, m_0)$ as we saw before.

Since A^M is $P \times_G TM$-valued, we can write

$$\left\langle A^M\left(\frac{\mathrm{d}pm}{\mathrm{d}t}\right), p_\alpha\right\rangle = \langle p\dot{m} + p(A(\dot{p})m), p\alpha\rangle = \langle \dot{m} + vm, \alpha\rangle$$

$$= \theta(\alpha)(\dot{\alpha}) + \langle \mathbf{J}(\alpha), v\rangle.$$

Here $vm = v_m(m)$ stands for the infinitesimal generator of $v = A(p)$ calculated at m, θ is the canonical 1-form on T^*M and $\mathbf{J}$ is the momentum mapping of the action of G on T^*M.

Let $L : TP \to \mathbf{R}$ be a given Lagrangian. Define $L^M : TP \times TT^*M \to \mathbf{R}$ by

$$L^M(p, \dot{p}; m, \alpha, \dot{m}, \dot{\alpha}) = L(p, \dot{p}) + \left\langle A^M\left(\frac{\mathrm{d}(pm)}{\mathrm{d}t}\right), p_\alpha\right\rangle.$$

Using previous formulas we have

$$\begin{aligned} L^M(p, \dot{p}; m, \alpha, \dot{m}, \dot{\alpha}) &= L(p, \dot{p}) + \langle p\dot{m} + p \cdot vm, p\alpha\rangle \\ &= L(p, \dot{p}) + \langle \dot{m} + vm, \alpha\rangle \\ &= L(p, \dot{p}) + \theta(m, \alpha)(\dot{m}, \dot{\alpha}) + \langle \mathbf{J}(m, \alpha), v\rangle. \end{aligned}$$

THEOREM 3.3. *Let $p_0, p_1 \in P$ and $m_0 \in U$ be given. Then the following assertions are equivalent.*

(i) *The curve $p \in \Omega(P; p_0, p_1)$ is a critical point of the functional $S : \Omega(P; p_0, p_1) \to \mathbf{R}$ defined by*

$$S(p) = \int_{t_0}^{t_1} L(p, \dot{p})\,\mathrm{d}t$$

(ii) *There is a curve $(m, \alpha) \in \Omega(T^*U; m_0, m_1(p))$ such that the curve $(p; m, \alpha)$ is a critical point of the functional $S^M : \Omega(P \times T^*U; p_0, p_1, m_0) \to \mathbf{R}$ defined by*

$$S^M(p, m, \alpha) = \int_{t_0}^{t_1} L^M(p, \dot{p}, m, \alpha, \dot{m}, \dot{\alpha})\, dt.$$

Proof. This will be a typical Lagrange multiplier argument, based on the following version of the

Lagrange Multiplier Theorem: *Let E be a manifold, $h : E \to \mathbf{R}$ a function and $E' \subset E$ a constraint submanifold defined by an equation $\varphi = 0$, where $\varphi : E \to F$, F a vector space and 0 is a regular value of φ. Then, for $e_0 \in E'$, the following assertions are equivalent:*

(i) *e_0 is a critical point of h restricted to E'.*

(ii) *There exists $\alpha_0 \in F^*$ such that (e_0, α_0) is a critical point of the function $\mathscr{H}(e, \alpha) = h(e) + \langle \varphi(e), \alpha \rangle$.*

In our case, we take

$$E := \Omega(P \times M; p_0, p_1, m_0) \text{ and } h(p, m) := \int_{t_0}^{t_1} L(p, \dot{p})\, dt =: S(p)$$

which, accidentally, is independent of m, and

$$E' := \Omega^H(P \times M, p_0, p_1, m_0).$$

According to Corollary 3.2 we can write $E' \cong \Omega(P; p_0, p_1)$. By working in a local chart of U we can assume $TU = U \times H$ where H is a vector space isomorphic to the tangent space at a point of U, and therefore $T^*U = U \times H^*$. Define $F = \Omega(H) = \{a : [t_0, t_1] \to H\}$ and $F^* = \Omega(H^*) = \{\alpha : [t_0, t_1] \to H\}$. Define the pairing

$$\langle\langle a, \alpha \rangle\rangle = \int_{t_0}^{t_1} \langle a(t), \alpha(t) \rangle\, dt$$

where $\langle a(t), \alpha(t) \rangle$ is the canonical pairing between H and H^*. Define the constraint function $\varphi : E \to F$ by $\varphi(p, m) = \dot{m} + vm$. Recall that $\dot{m}$ being the derivative of $m \in \Omega(H)$ can also be interpreted as a curve belonging to $\Omega(H)$, since h is a vector space. Also recall that $v(t)m(t)$ is, for each t, the infinitesimal generator of $v(t)$ at $m(t)$, which can be identified with an element of the vector space H. Thus it makes sense to form the pairing

$$\langle\langle \dot{m} + vm, \alpha \rangle\rangle = \int_{t_0}^{t_1} \langle m + vm, \alpha \rangle\, dt$$

as before.

From the previous remarks we conclude that we can take $\mathscr{H}(p; m, \alpha) = S^M(p; m, \alpha)$ and therefore, the assertion of our theorem follows from the Lagrange Multipler Theorem. [However, it should be mentioned that the usual versions of Lagrange Multipler theorem that appear in the literature are not strong enough to be applied directly in the way we described before, to many cases of physical interest. We claim, however, that the result remains valid in all those examples, and the proof can be performed by first localizing and then applying the same method as we did in an earlier paper (see Cendra and Marsden [1987]). We omit the details here]. ■

4. EQUATIONS OF MOTION

We will call $L : TP \to \mathbf{R}$ an *invariant Lagrangian* if the following condition is satisfied: $L(pg, \dot{p}g) = L(p, \dot{p})$ for all $(p, \dot{p}) \in T_pP$ and $g \in G$, where the notation «$(pg, \dot{p}g)$» stands for the tangent lifting of the action of G on P. We assume, from now on, that L is an invariant Lagrangian.

First of all, recall that any given connection A on P gives rise to a decomposition of TP into a «horizontal» part and a «vertical» part. This is related to Montomery [1986], where the cotangent version of these ideas have been studied in connection with the Hamiltonian (rather than Lagrangian) description of a system on a principal bundle. Examples such as free boundary fluids and particles in Yang-Mills field which were studied in Montgomery's work, can also be studied from the Lagrangian point of view, using the methods of the present article.

The decomposition of TP is given by an isomorphism

$$\phi : TP \to T^HP \times g$$

defined by $\phi(p, \dot{p}) = (p, \dot{p}^H, v) = : (H(p, \dot{p}), A(p, \dot{p}))$. The notation $H(p, \dot{p}) = (p, \dot{p}^H)$ stands for the horizontal component of $(p, \dot{p}) \in T_pP$. Thus $H : TP \to T^HP$ is, by definition, an onto map. Observe that H can be expressed in terms of A as follows: $H(p, \dot{p}) = (p, \dot{p}) - pA(p, \dot{p})$. The inverse of ϕ is given by $\phi^{-1}(p, \dot{p}^H, v) = (p, \dot{p}^H + pv)$, where, as we usually do in this paper, $pv = v_p(p)$ is the infinitesimal generator of v calculated at $p \in P$. The action of G on TP, which is the tangent lifting of the action of G on P, becomes

$$(p, \dot{p}^H, v)g = (pg, \dot{p}^H g, Ad_{g^{-1}} v).$$

With some abuse of notation we can write $L = L \circ \phi$, in other words, we shall write $L(p, \dot{p}) = L(H(p, \dot{p}), A(p, \dot{p}))$. Thus invariance of L is expressed by

$$L(p, \dot{p}^H, v) = L(pg, \dot{p}^H g, Ad_{g^{-1}} v).$$

Since L is invariant, it defines a function on $T^HP \times_G g$, where the action of G on g is the adjoint action. The main purpose of this section is to get equations of motion on $T^HP \times_G g \times T^*M$ from the variational principle described in the previous section. The curvature of the connection A will appear in a natural way as a term representing a field strength, which seems natural in view of the case of a particle in Yang-Mills field. This will be done using a description of $T^HP \times_G g$ as a vector bundle over TB with fiber isomorphic to g, which resembles a similar description for the cotangent case given in Montgomery [1986]. A local trivialization $X \times G$ of P gives rise to a local trivialization $TX \times g$ of $T^HP \times_G g$ and therefore our lagrangian L^M, locally, is a function $L_M : TX \times g \times T^*M \to \mathbb{R}$ and the functional S^M becomes, locally, a function

$$S^M : \Omega(X; x_0, x_1) \times \Omega(g) \times \Omega(T^*M; m_0, m_1(p)) \to \mathbb{R}.$$

The curve p that appears in $m_1(p)$ is determined by the curve (x, v) in $\Omega(X; x_0, x_1) \times \Omega(g)$, as follows: Solve $v = A(x, \dot{x}, g, \dot{g})$, $g(t_0) = e$, which gives a solution $g(t)$. Find the horizontal lift $x^h(t)$ of $x(t)$ such that $x^h(t_0) = p_0$, which is a curve in $\Omega(P; p_0, x_1)$. Then set $p(t) = x^h(t)g(t)$. The function m_1 was defined before. In other words, a typical element of

$$\Omega(X; x_0, x_1) \times \Omega(g) \times \Omega(T^*M; m_0, m_1(p))$$

can be described as follows. Choose

$$(x, v) \in \Omega(X; x_0, x_1) \times \Omega(g)$$

arbitrarily. Then find p and $m_1(p)$ as we did before. Finally choose a curve $(m(t), \alpha(t))$ on $\Omega(T^*M; m_0, m_1(p))$, i.e. a curve such that $m(t_0) = m_0$, $m(t_1) = = m_1(p)$ with $\alpha(t)$ arbitrary. Then (x, v, m, α) is a typical element of that space of curves. The idea is now to apply the usual variational techniques and notice that arbitrary variations of the curve x with fixed endpoints, arbitrary variations of v and arbitrary variations of (m, α) with conditions $m(t_0) = m_0$, $m(t_1) = m_1(p)$ are allowed. We will postpone giving a detailed description of the equations that we get by this procedure until the end of this section.

The next step consists of finding critical points of the functional S^M, which according to Theorem 3.3 is equivalent to finding critical points of the functional S (which in turn, is the typical functional of the Hamilton Variational principle). First of all let us introduce some notation. Recall the decomposition (see §3) $p(t) = p^h(t)g^p(t)$, where $p^h(t)$ is horizontal and $p^h(t_0) = p(t_0) = p_0$. Then we have $A(p, \dot{p}) = A(p^h g^p, \dot{p}^h g^p + p^h \dot{g}^p) = A(p^h g^p, p^h g^p (g^p)^{-1} \dot{g}^p) = v$, where $v = (g^p)^{-1} \dot{g}^p \in \Omega(g)$.

It is useful to keep in mind that, since for given $p_0 \in P$ and $p \in \Omega(P; p_0)$, the above decomposition is unique, so we have an isomorphism $\Omega(P; p_0) \to \Omega^H(P, p_0) \times$

$\times\ \Omega(g)$ given by $p \mapsto (p^h, A(p, \dot{p}))$ where $\Omega^H(P; p_0) = \{p \in \Omega(P; p_0) \mid p$ is horizontal$\}$. The inverse is constructed as follows: given $v \in \Omega(g)$, find $g \in \Omega(G; e)$ such that $\dot{g} = gv$. (This amounts to solving a time dependent ordinary differential equation on G). Thus $p = p^h g$ is the inverse image of the curve (p^h, v). Also observe that $\Omega^H(P; p_0) \cong \Omega(B; x_0)$, since the map $p^h \mapsto \Omega(\pi)(p^h) := \pi \circ p^h$ is an isomorphism.

Now we apply variations to curves in $\Omega P \times T^*U, p_0, p_1, m_0)$ to get equations of motion, according to Theorem 3.3. From now on « · » will be used to denote derivatives with respect to the variable t only, and we will sometimes use a notation of type δp, δu, etc. to denote derivatives with respect to the parameter λ.

An arbitrary vertical variation of the curve p, compatible with the endpoint conditions $p(t_0) = p_0$ and $\pi(p(t_1)) = x_1$, can be represented by $p_\lambda = pg_\lambda$, where $g_\lambda \in \Omega(G; e)$ is arbitary. Thus $v'_\lambda = (g_\lambda)^{-1}\dot{g}_\lambda$ is an arbitrary variation of $v: = g^{-1}\dot{g}$ in $\Omega(g)$. We have

$$L\left(H\left(\frac{\mathrm{d}pg_\lambda}{\mathrm{d}t}\right), A\left(\frac{\mathrm{d}pg_\lambda}{\mathrm{d}t}\right)\right) = L(H(\dot{p}g_\lambda + p\dot{g}_\lambda) + A(\dot{p}g_\lambda + p\dot{g}_\lambda))$$

$$= L((pg_\lambda + \dot{p}^H g_\lambda), (Ad_{g_\lambda^{-1}} A(p, \dot{p}) + v'_\lambda)).$$

Because of the invariance of L, the last expression becomes $L(p, \dot{p}^H, v + u_\lambda)$ where $v = A(p, \dot{p})$ and $u_\lambda = \dot{g}_\lambda (g_\lambda)^{-1}$. Observe that u_λ also represents an arbitrary variation of v. Thus we have

$$\left.\frac{\partial L(p_\lambda, \dot{p}_\lambda^H, v_\lambda)}{\partial \lambda}\right|_{\lambda=0} = \left.\frac{\partial L(p_\lambda, \dot{p}_\lambda)}{\partial \lambda}\right|_{\lambda=0}$$

$$= \frac{\partial L}{\partial v}(p, \dot{p}^H, v)\left.\left\langle\frac{\partial u_\lambda}{\partial \lambda}\right\rangle\right|_{\lambda=0} = \frac{\partial L}{\partial v}(p, \dot{p}^H, v)\langle \delta v\rangle$$

where $\delta v = \left.\dfrac{\partial u_\lambda}{\partial \lambda}\right|_{\lambda=0}$.

Variations (p_λ, m, α), where $p_\lambda = pg_\lambda$ are compatible with the imposition of the constraint $\Omega(P \times T^*U; p_0, p_1, m_0)$. This is because, according to the definition of m_1,

$$m_1(p_\lambda) = i_p^{-1}[p_\lambda{}^h(t_1) m_0] = i_p^{-1}[p^h(t_0) m_0]$$

which does not depend on λ and therefore we can assume that $(m_\lambda, \alpha_\lambda)$ is independent of λ. Then we have

$$\frac{\partial}{\partial \lambda}\langle \dot{m} + v_\lambda m, \alpha\rangle|_\lambda = J(m, \alpha)\langle \delta v\rangle$$

where

$$\delta v = \left.\frac{\partial v_\lambda}{\partial \lambda}\right|_{\lambda=0}.$$

Since $L^M(p, \dot{p}, m, \dot{m}, \alpha, \dot{\alpha}) = L(p_1, \dot{p}_\lambda{}^H, v) + \langle \dot{m} + v_\lambda m, \alpha\rangle$. We can conclude that vertical variations of the curve p, lead to the equation

$$\frac{\partial L}{\partial v}(p, \dot{p}^H, v) = -\mathbf{J}(m, \alpha). \qquad (\delta p)^v$$

Let p_λ be a horizontal variation of the curve p, with fixed endpoints, i.e., for each t, $p_\lambda(t) \equiv p(\lambda, t)$ is a horizontal curve and $p_\lambda(t_i) = p(t_i)$, $i = 0, 1$. Choose, for each λ, the curve m_λ so that $p_\lambda m_\lambda$ is horizontal. According to §3 this can be always achieved in a unique way, by taking $m_\lambda = m^{p_\lambda}$. This implies that $L^M(p_\lambda, \dot{p}_\lambda, m, \dot{m}, \alpha, \dot{\alpha}) = L(p_\lambda, \dot{p}_\lambda)$ and therefore we should only study

$$\left.\frac{\partial}{\partial \lambda}\int_{t_0}^{t_1} L(p_\lambda, \dot{p}_\lambda)\, dt\right|_{\lambda=0}.$$

By the usual integration by parts argument, we get the equation

$$\left(\frac{\partial L}{\partial p}(p, \dot{p}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{p}}(p, \dot{p})\right)(\delta p)^H = 0 \qquad (\delta p)^H$$

for all horizontal vectors

$$\left.\frac{\partial p}{\partial \lambda}\right|_{\lambda=0} = (\delta p)^H \text{ at } p \in P.$$

While the equations $(\delta p)^H$ appear formally like Euler-Lagrange equations, they are not unless the distribution of horizontal planes in P is integrable, i.e., the connection A is a trivial connection.

One of our purposes is to give another expression for equation $(\delta p)^H$, which involves the curvature of the connection A. This will be done at the end of this section.

Next we consider variations $(m_\lambda, \alpha_\lambda)$ of the curve (m, α) satisfying the fixed endpoint condition $m(t_0) = m_0$, $m(t_1) = m_1(p)$, so variations $(p, m_\lambda, \alpha_\lambda)$ are compatible with the constraint $\Omega(P \times T^*U; p_0, m_1, m_0)$. To simplify notation, we will write $\gamma = (m, \alpha)$ and we will assume that the variation $\gamma_\lambda = (m_\lambda, \alpha_\lambda)$ satisfies the fixed endpoint condition $\delta\gamma(t_i) = 0$, $i = 0, 1$, which is compatible with the previous constraint. Since θ is the canonical 1-form on T^*M and

$\omega = -\mathbf{d}\theta$ is the canonical symplectic form on T^*M we can check (see for instance Cendra and Marsden [1987]) that

$$\delta \int_{t_0}^{t_1} \theta(\dot{\gamma})\, \mathrm{d}t = \int_{t_0}^{t_1} \omega(\dot{\gamma}, \delta\gamma)\, \mathrm{d}t\,. \qquad (\delta\gamma)$$

As we explained before, $\delta\gamma$ stands for the derivative of γ with respect to λ, and $\dot{\gamma}$ is the derivative of γ with respect to t. From the definition of the momentum mapping, we know that $\mathbf{J}(v)(\gamma) = \langle \mathbf{J}(\gamma), v\rangle$ is the Hamiltonian of the infinitesimal generator of the action of G on T^*M corresponding to $v \in \mathfrak{g}$. Thus we have

$$\delta \int_{t_0}^{t_1} \mathbf{J}(\gamma)(v)\, \mathrm{d}t = \int_{t_0}^{t_1} [\mathrm{d}\mathbf{J}(v)(\gamma)(\delta\gamma)]\, \mathrm{d}t$$

$$= \int_{t_0}^{t_1} \omega(v\gamma, \delta\gamma)\, \mathrm{d}t\,.$$

Since ω is nondegenerate, equation $(\delta\gamma)$ is equivalent to

$$\dot{\gamma} + v\gamma = 0, \qquad (\delta\gamma)'$$

which represents the «Lin Constraint». Let us summarize the equations of motion that we have obtained so far for convenience:

$$\frac{\partial L}{\partial v}(p, \dot{p}^H, v) = -\mathbf{J}(\gamma) \qquad (\delta p)^v$$

$$\left(\frac{\partial L}{\partial p} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{p}}\right)((\delta p)^H) = 0 \qquad (\delta p)^H$$

for all horizontal vectors $(\delta p)^H$ at $p \in P$ and

$$\dot{\gamma} + v\gamma = 0. \qquad (\delta\gamma)$$

Now we give an interesting expression for equation $(\delta p)^H$ which involves the curvature of the connection A. This will be done by first choosing a local trivialization of P, say $X \times G$ where X is an open set contained in B. Thus we have an isomorphism: $T(X \times G) \cong TX \times G \times \mathfrak{g}$ given by $(x, \dot{x}, g, \dot{g}) \mapsto (x, \dot{x}, g, A(x, \dot{x}, g, \dot{g})) := (x, \dot{x}, g, v)$. [**Warning**: vectors of the form $(x, \dot{x}, g, 0)$ are not necessarily horizontal, while vectors $(x, 0, g, \dot{g})$ are always vertical, with respect to the connection A]. The action of G on $TX \times G \times \mathfrak{g}$ induced by this trivialization now becomes $(x, \dot{x}, g, v)h = (x, \dot{x}, gh, Ad_{h^{-1}}v)$. Therefore we can write

$TX \times G \times g/G \cong TX \times g$, and the canonical projection $TX \times G \times g \mapsto TX \times g$ is given by $(x, \dot{x}, g, v) \mapsto (x, \dot{x}, Ad_g v)$. Since L is invariant, it induces a function L' on $TX \times g$ given by

$$L'(x, \dot{x}, v) = L(x, \dot{x}, e, v) = L(x, \dot{x}, g, Ad_{g-1} v) \text{ where } v = A(p, \dot{p}).$$

Now we are ready to study equations $(\delta p)^H$. As before, choose variations $p_\lambda(t)$ of the curve $p(t) = p_0(t)$ such that, for each t, $p_\lambda(t)$ is a horizontal curve with $p_\lambda(t_i)$ is fixed for $i = 0, 1$. Choose $m_\lambda = m^{p_\lambda}$, so that

$$A^M\left(\frac{\mathrm{d}(p_\lambda m_\lambda)}{\mathrm{d}t}\right) = 0, \text{ for all } t, \lambda.$$

Denoting $v_\lambda(t) = A(p_\lambda(t), \dot{p}_\lambda(t))$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} S^M(p_\lambda; m_\lambda, \alpha_\lambda)\Bigg|_{\lambda=0} = \frac{\mathrm{d}}{\mathrm{d}\lambda} \int_{t_0}^{t_1} L'(x_\lambda, \dot{x}_\lambda, Ad_{g_\lambda} A(p_\lambda, \dot{p}_\lambda))\, \mathrm{d}t \Bigg|_{\lambda=0}$$

$$= \int_{t_0}^{t_1} \left[\left(\frac{\partial L'}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L'}{\partial \dot{x}}\right)\langle\delta x\rangle + \left\langle\frac{\partial L'}{\partial u} \frac{\partial}{\partial \lambda} Ad_{g_\lambda} A(p_\lambda, \dot{p}_\lambda)\right\rangle\right] \mathrm{d}t \Bigg|_{\lambda=0}$$

Since $A(p, \dot{p})$ is linear in $\dot{p}$, write $A(p, \dot{p}) = A(p)\dot{p}$. Then we have

$$\int_{t_0}^{t_1} \left\langle\frac{\partial L'}{\partial u} \frac{\partial}{\partial \lambda} Ad_{g_\lambda} A(p_\lambda)\dot{p}_\lambda\right\rangle\Bigg|_{\lambda=0} \mathrm{d}t = \int_{t_0}^{t_1} \frac{\partial L'}{\partial u} \left\langle Ad_{g_0} \frac{\partial A(p_\lambda)\dot{p}_\lambda}{\partial \lambda}\right\rangle\Bigg|_{\lambda=0} \mathrm{d}t$$

$$+ \int_{t_0}^{t_1} \left\langle\frac{\partial L'}{\partial u} \frac{\partial Ad_{g_\lambda}}{\partial \lambda} A(p_0)\dot{p}_0\right\rangle\Bigg|_{\lambda=0} \mathrm{d}t = : C + D. \qquad (CD)$$

Set $g_0 = g$, $p_0 = p$ for simplicity. Then the integral C defined by the equation (CD) is given by

$$C = \int_{t_0}^{t_1} \frac{\partial L'}{\partial u} \circ Ad_g \left(\frac{\partial A(p_\lambda)p_\lambda}{\partial \lambda}\right)\Bigg|_{\lambda=0} \mathrm{d}t$$

$$= \int_{t_0}^{t_1} \frac{\partial L'}{\partial u} \circ Ad_g \left(\frac{\partial A}{\partial p}\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right) + A(p_\lambda)\frac{\partial^2 p}{\partial \lambda \partial t}\right)\Bigg|_{\lambda=0} \mathrm{d}t$$

$$= \int_{t_0}^{t_1} \frac{\partial L'}{\partial u} \circ Ad_g\left(\frac{\partial A}{\partial p}\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right) - \frac{\partial A}{\partial p}\left(\frac{\partial p}{\partial t}, \frac{\partial p}{\partial \lambda}\right)\right)$$

$$- \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial u} \circ Ad_g\right) A(p_\lambda)\, \frac{\partial p}{\partial \lambda}\bigg|_{\lambda=0} \mathrm{d}t,$$

where the last equality comes from integrating by parts. Since the derivative of p with respect to λ is horizontal we have

$$A(p_\lambda)\, \frac{\partial p}{\partial \lambda} = 0.$$

This and the Cartan Structure equation gives

$$\frac{\partial A}{\partial p}\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right) - \frac{\partial A}{\partial p}\left(\frac{\partial p}{\partial t}, \frac{\partial p}{\partial \lambda}\right) = \mathrm{d}A\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right) = \frac{1}{2}\left[A\left(\frac{\partial p}{\partial \lambda}\right), A\left(\frac{\partial p}{\partial t}\right)\right]$$

$$+ \Omega\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right) = \Omega\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right)$$

where Ω is the curvature form of the connection A. Thus if we set

$$c = \frac{\partial L'}{\partial u}\left\langle Ad_g \Omega\left(\frac{\partial p}{\partial \lambda}, \frac{\partial p}{\partial t}\right)\right\rangle\bigg|_{\lambda=0}.$$

We have

$$C = \int_{t_0}^{t_1} c\, \mathrm{d}t.$$

On the other hand, if we set

$$w'(t) = [g(t)]^{-1}\, \frac{\partial g_\lambda(t)}{\partial \lambda}\bigg|_{\lambda=0'}$$

we can readily check that

$$d := \frac{\partial L'}{\partial u}\left\langle \frac{\partial}{\partial \lambda}\, Ad_{g_\lambda} A(p)\dot{p}\right\rangle\bigg|_{\lambda=0} = \frac{\partial L'}{\partial u}\left\langle \frac{\partial}{\partial \lambda}\, Ad_g[w', v]\right\rangle$$

and therefore

$$D = \int_{t_0}^{t_1} d \, \mathrm{d}t.$$

We can give c and d a more symmetric expression as follows. Since $p(\lambda, t) = (x(\lambda, t), g(\lambda, t))$, we can write $p = qh$ where $h(\lambda, t) = [g(t)]^{-1} g(\lambda, t)$, giving $h(0, t) = e$ and $q(\lambda, t) = (x(\lambda, t), g(t))$. Thus

$$A\left(\frac{\partial q}{\partial \lambda}\right)\bigg|_{\lambda=0} = A\left(\frac{\partial p h^{-1}}{\partial \lambda}\right)\bigg|_{\lambda=0} = A \frac{\partial q}{\partial \lambda} + A(p(-w')) = -w'$$

because $-w' = \partial h^{-1}/\partial\lambda\,|_{\lambda=0}$, and using the fact that $\mathrm{d}p/\mathrm{d}\lambda$ is horizontal Since $p(0, t) = q(0, t)$, we have

$$d = \frac{\partial L'}{\partial u} Ad_g\left[A\left(\frac{\partial q}{\partial t}\right), A\left(\frac{\partial q}{\partial \lambda}\right)\right].$$

We can also check that

$$c = \frac{\partial L'}{\partial \mu} Ad_g \Omega\left(\frac{\partial q}{\partial t} \quad \frac{\partial q}{\partial \lambda}\right).$$

Finally, we can show that c and d are invariant under the action of G and therefore they define functions $c'(x, \dot{x}, \delta x, u)$ and $d'(x, \dot{x}, \delta x, u)$, where as usual δx stands for the derivative of $x(\lambda, t)$ with respect to λ. The usual variational techniques, starting with equation (CD), equation $(\delta p)^H$ becomes

$$\frac{\partial L'}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L'}{\partial x} = -(c'(x, \dot{x}, \bullet, u) + d'(x, \dot{x}, \bullet, u)).$$

To find an explicit expression for the equations of notion, we first calculate c', d'. Introduce the notion $s := \dot{g}g^{-1}$ for convenience. We can check that for a given element $s \in \mathfrak{g}$ we have $s = A(x, 0, e, s)$ with $A'(x, \dot{x}) := A(x, \dot{x}, e, 0)$. Thus,

$$\begin{aligned} v = A(x, \dot{x}, g, \dot{g}) &= A(x, \dot{x}, g, 0) + A(x, 0, g, \dot{g}) = A(x, \dot{x}, g, 0) + A(x, 0, g, sg) \\ &= Ad_{g^{-1}}A(x, \dot{x}, e, 0) + Ad_{g^{-1}}\dot{g}g^{-1} = Ad_{g^{-1}}A^1(x, \dot{x}) + Ad_{g^{-1}}s. \end{aligned}$$

Also, since $u = Ad_g v$, we have $u = A'(x, \dot{x}) + s$. Then we have

$$\begin{aligned} Ad_g\left[A\left(\frac{\partial q}{\partial t}\right), A\left(\frac{\partial q}{\partial \lambda}\right)\right] &= \left[\left[A\left(\frac{\partial q}{\partial t} g^{-1}\right), A\left(\frac{\partial q}{\partial t} g^{-1}\right)\right]\right] \\ &= \left[A\left(\left(x, \frac{\partial x}{\partial t}, g, \frac{\partial g}{\partial t}\right)g^{-1}, A\left(\left(x, \frac{\partial x}{\partial \lambda}, g, 0\right)g^{-1}\right)\right]\right] \end{aligned}$$

$$= \left[A\left(x, \frac{\partial x}{\partial t}, e, \frac{\partial g}{\partial t} g^{-1}\right), A\left(x, \frac{\partial x}{\partial \lambda}, e, 0\right) \right]$$

$$= \left[A\left(x, \frac{\partial x}{\partial t}, e, 0\right) + A(x, 0, e, s), A\left(x, \frac{\partial x}{\partial \lambda}, e, 0\right) \right]$$

$$= \left[u, A'\left(x, \frac{\partial x}{\partial \lambda}\right) \right].$$

Therefore,

$$d'(x, \dot{x}, \delta x, u) = \frac{\partial L'}{\partial u} [u, A'(x, \delta x)].$$

Now define $\Omega'(x, \dot{x}, \delta x) := \Omega((x, \dot{x}, e, 0), (x, \delta x, e, 0))$. Hence

$$Ad_g \Omega\left(\frac{\partial q}{\partial \lambda}, \frac{\partial q}{\partial t}\right)\bigg|_{\lambda=0} = \Omega\left(\frac{\partial q}{\partial \lambda} g^{-1}, \frac{\partial q}{\partial t} g^{-1}\right)\bigg|_{\lambda=0}$$

$$= \Omega\left(\left(x, \frac{\partial x}{\partial \lambda} g, \frac{\partial g}{\partial \lambda}\right), \left(x, \frac{\partial x}{\partial t}, g, \frac{\partial g}{\partial t}\right)\right)\bigg|_{\lambda=0}$$

$$= \Omega'\left(\left(x, \frac{\partial x}{\partial \lambda} e, \frac{\partial g}{\partial \lambda} g^{-1}\right), \left(x, \frac{\partial x}{\partial t}, e, \frac{\partial g}{\partial t} g^{-1}\right)\right)$$

$$= \Omega'\left(x, \frac{\partial x}{\partial \lambda}, \frac{\partial g}{\partial \lambda}\right)$$

because Ω is a tensorial form and $(x, 0, e, (\partial g/\partial \lambda) g^{-1})$ and $(x, 0, e, (\partial g/\partial t) g^{-1})$ are vertical vectors. Then we have

$$c'(x, \dot{x}, \delta x, u) = \frac{\partial L'}{\partial u} \Omega'(x, \delta s, \dot{x}).$$

So far we have proved that equation $(\delta p)^H$ becomes

$$\frac{\partial L'}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L'}{\partial \dot{x}} = \frac{\partial L'}{\partial u} (\Omega'(x, \dot{x}, \bullet) + [A(x, \bullet), u]).$$

Equation $(\delta\gamma)$ can be written as follows

$$\dot{\gamma} + v\gamma = 0 \Longleftrightarrow g\dot{\gamma} + \dot{g}\gamma - \dot{g}g^{-1}g\gamma + gvg^{-1}g\gamma = 0.$$

Setting $\beta = g\gamma$, and using the preceding equivalence, we have

$$\dot{\gamma} + v\gamma = 0 \Longleftrightarrow \dot{\beta} + u\beta - s\beta = 0 \Longleftrightarrow \dot{\beta} + A'(x, \dot{x})\beta = 0.$$

In other words, equation $(\delta\gamma)$ is equivalent to $\dot{\beta} + A'(x, \dot{x})\beta = 0$. Finally equation $(\delta p)^v$ was

$$\frac{\partial L}{\partial v} = -\mathbf{J}(\gamma).$$

Since $u = Ad_g v$, we have

$$\frac{\partial L}{\partial v} = \frac{\partial L'}{\partial u} \circ Ad,$$

Consequently $(\delta p)^v$ is equivalent to

$$\frac{\partial L'}{\partial u} = -\mathbf{J}(\beta).$$

Collecting these results, our main equations of motion become

$$\frac{\partial L'}{\partial u} = -\mathbf{J}(\beta) \qquad (\delta p)^v$$

$$\frac{\partial L'}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L'}{\partial \dot{x}} = \frac{\partial L'}{\partial u}(\Omega'(x, \dot{x}, \bullet) + [A'(x, \dot{x}), u]) \qquad (\delta p)^H$$

$$\dot{\beta} + A'(x, \dot{x})\beta = 0 \qquad (\delta\gamma)$$

where

$$A'(x, \dot{x}) = A(x, \dot{x}, e, 0),$$

$$\Omega'(x, \dot{x}, \delta x) = \Omega((x, \dot{x}, e, 0), (x, \delta x, e, 0)),$$

and

$$L'(x, \dot{x}, u) = L(x, \dot{x}, e, u).$$

If we find a solution curve say $(x(t), u(t), \beta(t))$ to the main equations, then we can also find $g(t)$ by solving the equation

$$Ad_g A(x, \dot{x}, g, \dot{g}) = u.$$

In this way we can *reconstruct* the motion on P using $p(t) = (x(t), g(t))$ and $\gamma(t) = \beta(t)[g(t)]^{-1}$. Once the local trivialization $X \times G$ of P and the connection A have been chosen, then we can write A', Ω', L' and the equations of motion, without any further calculation, by using the above expressions.

Remarks 1. If the Lagrangian L' has the stronger invariance property given by the condition $L'(x, \dot{x}, u) = L'(x, \dot{x}, Ad_g u)$ for all $g \in G$, (this happens in the example of particles in a Yang-Mills field as a consequence of the bi-invariance of the metric defined on the group), then we have

$$d = \frac{dL}{\partial u}\left\langle \frac{\partial}{\partial \lambda} Ad_{g_\lambda} A(p)\dot{p} \right\rangle \Bigg|_{\lambda=0} = \frac{d}{d\lambda} L'(x, \dot{x}, Ad_{g_\lambda} A(p)\dot{p}) \Bigg|_{\lambda=0} = 0.$$

In this case the equations of motion become

$$\frac{dL'}{\partial u} = -\mathbf{J}(B)$$

$$\frac{\partial L'}{\partial x} - \frac{d}{dt}\frac{\partial L'}{\partial \dot{x}} = \frac{\partial L'}{\partial u}\Omega'(x, \dot{x}, \bullet)$$

$$\beta + A'(x, \dot{x})\beta = 0.$$

2. Let $L : TP \to \mathbf{R}$ be an invariant Lagrangian (not necessarily bi-invariant as it was in remark 1). Choose a local trivialization $X \times G$ of P as before. We have an isomorphism $T(X \times G)/G \cong TX \times \mathfrak{g}$ given by $[(x, \dot{x}, g, \dot{g})] \mapsto (x, \dot{x}, v)$ where $v = \dot{g}g^{-1}$. We can check that this is well defined, showing that we do not really need a connection to establish an isomorphism as before, once we have chosen a trivialization. Since L is an invariant Lagrangian, it induces a «reduced» Lagrangian on $TX \times \mathfrak{g}$. This is exactly the situation considered in Cendra and Marsden [1987]. Then we can introduce Clebsch variables as we did in that paper. Thus we get an alternative approach to the question of dividing by the symmetry of the system, which does not involve a connection. The possibility of this double approach to a given system with symmetry has a cotangent counterpart (see Montgomery [1986]).

5. YANG-MILLS SPACES

We now describe the motion of a particle in a Yang-Mills field as an illustration of the ideas previously discussed about variational principles in principal fiber bundles. The configuration space for a particle moving in a Yang-Mills field F is a manifold B, the base space of a principal fiber bundle $\pi : P \to B$, called the gauge configuration space and F is the curvature of a connection A on P. The Lagrangian is the kinetic energy Lagrangian corresponding to a metric K on the total space P. This metric is constructed by glueing together a Riemannian metric g on the base space B and a bi-invariant metric λ on the group manifold G using the connection A. The metric K is defined by the formula

$$K_p(\dot{p}, \dot{q}) = g_{\pi(p)}(T_p\pi(\dot{p}), T_p\pi(\dot{q})) + \lambda(A(\dot{p}), A(\dot{q})).$$

for every pair of tangent vectors $\dot{p}$, $\dot{q} \in T_pP$. The projection of the geodesic motion on P onto B with respect to the metric K, gives the solution to Wong's equations (see Montgomery [1984]). This Kaluza-Klein approach was first devised in Kerner [1968] but its canonical counterpart was not fully understood until very recently (Sternberg [1977], Weinstein [1978], Montgomery [1984]). The following paragraph provides a brief account of the Hamiltonian description of the theory.

The kinetic energy Lagrangian

$$L(p, \dot{p}) = \frac{1}{2} K_p(\dot{p}, \dot{p})$$

can be used to identify TP with T^*P. The Hamilotnian induced on T^*P in this way will be denoted by H, and is given by

$$H(p, \alpha) = \frac{1}{2} K_p^{\flat}(\alpha, \alpha)$$

for every $\alpha \in T_p^*P$. Hence «$\flat$» stands for the operation of lowering indices. This Hamiltonian is obviously G-invariant with respect to the cotangent lifting of the action of the Gauge group G. Using the symplectic reduction theorem we get a family of reduced Hamiltonian systems $((T^*P)_\mu, \Omega_\mu, H_\mu)$ where $(T^*P)_\mu$ denotes the reduced space obtained by taking the quotient of the level set of the momentum mapping $\mathbf{J} : T^*P \to g^*$ by G_μ, the isotropy group of the coadjoint action corresponding to $\mu \in g^*$. These spaces are the universal representation of the phase space of the system (Weinstein [1978]). The connection A allows us to identify $(T^*P)_\mu$ with the space $P^\# \times_G O_\mu$, where $P^\#$ denotes the pull back of P along the canonical projection $\tau_B^* : T^*B \to B$ and O_μ is the coadjoint orbit of G through $\mu \in g^*$. The symplectic structure on $P^\# \times_G O_\mu$ was described in Sternberg [1977] and it is just the projection of $P^\# \times_G O_\mu$ of the closed 2-form $\tilde{\Omega}_\mu = \omega + \omega_\mu + \mathbf{d}\langle \mathbf{J}, A\rangle$ defined on $P^\# \times_G O_\mu$, ω is the presymplectic form on $P^\#$ obtained by pulling-back the standard orbit symplectic structure on O_μ and $\theta_A = \langle \mathbf{J}, A\rangle$ is the couping 1-form on $P^\# \times O_\mu$. This form is projectable and induces the symplectic structure Ω_μ. The difference between H_μ and the Hamiltonian is a Casimir on g^* that does not affect the equations of motion.

If instead of doing the reduction by the group G in the Hamiltonian formalism we tried naively to reduce the Lagrangian system given by the metric K we would get the Lagrangian

$$L'(x, \dot{x}, v) = \frac{1}{2} g_x(\dot{x}, \dot{x}) + \frac{1}{2} \lambda(v, v)$$

where $(x, \dot{x}, v)$ denotes local coordinates in the bundle $TPG \cong TX \times g$, where $X \times G$ is a trivialization of P as we saw in §4. The naive variational principle, which allows arbitrary variations of curves x and v, does not provide the correct equations of motion. We could say that reduction and variational equations do not commute in a trivial way. However, we can write equations of motion as we have explained before, according to Remark 1 at the end of §4. Identifying vectors and covectors via the index lowering operation «♭» with respect to the metrics g on B and λ on G, we get

$$\frac{\partial L'}{\partial u} = -\mathbf{J}(\beta) = u^b$$

$$\nabla_{\dot{x}} \dot{x} = \lambda(u, \Omega'(x, \dot{x}, \bullet))$$

$$\dot{\beta} + A'(x, \dot{x})\beta = 0.$$

For the space of Clebsch potentials we have several choices, for example $M = G$, or M a vector space, as we have explained before. (See Balachandran et. al. [1985] for a description of this system along these lines).

REFERENCES

A.P. BALACHANDRAN, G. MARMO, B.S. SKAGERTAN and A. STERN [1985]: *Gauge Symmetries and Fiber Bundles*, Lecture Notes in Physics 188, Springer-Verlag.

H. CENDRA and J.E. MARSDEN [1987]: *Lin Constraints Clebsch Potentials and Variational Principles*, Physica **27D**, 63 - 89.

R. KERNER [1968]: *Generalization of the Kaluza-Klein Theory for an arbitrary non abelian Gauge Group*, Ann. Inst. Henri Poincaré, 9, 143 - 152.

S. KOBAYASHI and K. NOMIZU [1963]: *Foundations of Differential Geometry*, 2 vols. John Wiley.

R. MONTGOMERY [1984]: *Canonical Formulations of a Classical Particle in a Yang-Mills Field and Wong's Equations*, Lett. Math. Phys. **8**, 59 - 67.

R. MONTGOMERY [1986]: *The Bundle Picture in Mechanics*, Thesis, U.C. Berkeley.

W.A. POOR [1981]: *Differential Geometric Structures*, McGraw-Hill Co.

S. STERNBERG [1977]: *Minimal Coupling and the Symplectic Mechanics of a Classical Particle in the Presence of a Yang-Mills Field*, Proc. Nat. Acad. Sci. U.S.A. **74**, 5253 - 5254.

A. WEINSTEIN [1978]: *A Universal Phase Space for a Particle in a Yang-Mills Field*, Lett. Math. Phys. 2, 417 - 420.

Manuscript received: August 23, 1987.