

Park City Lectures on Mechanics, Dynamics, and Symmetry

Jerrold E. Marsden

July 17, 1997, this version: January 18, 1998

Contents

Preface	iv
1 Reduction for Mechanical Systems with Symmetry	1
2 Stability, Underwater Vehicle Dynamics and Phases	13
3 Systems with Rolling Constraints and Locomotion	31
4 Optimal Control and Stabilization of Balance Systems	43
5 Variational Integrators	53

Preface

In these five lectures, I cover selected items from the following topics:

1. Reduction theory for mechanical systems with symmetry,
2. Stability, bifurcation and underwater vehicle dynamics,
3. Systems with rolling constraints and locomotion,
4. Optimal control and stabilization of balance systems,
5. Variational integrators.

Each topic itself could be expanded into several lectures, but I limited myself to what I could reasonably explain in the allotted time. The hope is that the overview is informative enough so that the reader can understand the fundamental ideas and can intelligently choose from the literature for additional details on topics of interest.

Compatible with the theme of the PCI graduate school, I assume that the readers are familiar with the elements of geometric mechanics, including the basics of symplectic and Poisson geometry. The reader can find the needed background in, for example, Marsden and Ratiu [1998].

Acknowledgments. These lectures present a survey of *some* of the topics done in collaboration with many coauthors over the last few years. Specific reference is made directly in the text, but I am very grateful for their rewarding collaboration.

The organizers, staff and participants at the Park City Institute are to be congratulated for producing a stimulating and productive event. I would also like to thank the University of California, Berkeley, the California Institute of Technology and the National Science Foundation for helping support my research over the years.

Pasadena, CA
January, 1998

Jerrold E. Marsden

Lecture 1

Reduction for Mechanical Systems with Symmetry

In this first lecture, I will discuss a selection of topics from the basic principles of reduction theory for mechanical systems with symmetry. This is a rather large subject and many interesting aspects have to be eliminated or treated only briefly.

Reduction is of two sorts, Lagrangian and Hamiltonian. In each case one has a group of symmetries and one attempts to pass the structure at hand to an appropriate quotient space.

With Lagrangian reduction, the crucial object one wishes to reduce is *Hamilton's variation principle* for the Euler-Lagrange equations. On the other hand, with Hamiltonian reduction, the crucial objects to reduce are *symplectic and Poisson structures*. We begin by recalling Hamilton's principle.

The Euler-Lagrange equations and Hamilton's principle. We learn in mechanics that for a Lagrangian L defined on the tangent bundle TQ of a configuration manifold Q that the *Euler-Lagrange equations* for a curve $q(t) \in Q$; namely,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0$$

(in a coordinate chart on Q), may be regarded as a map (independent of coordinates) from the second order subbundle of TTQ to T^*Q . This is perhaps easiest to see through the equivalence of the Euler-Lagrange equations to *Hamilton's principle*:

$$\delta \mathfrak{S} = \delta \int L(q(t), \dot{q}(t)) dt = 0.$$

Here, \mathfrak{S} is the *action function*, the integral of the Lagrangian along the time derivative of a curve in Q with fixed endpoints, regarded as a function of that curve. Hamilton's principle states that *this function \mathfrak{S} on the space of curves has a critical point at a curve if and only if that curve satisfies the Euler-Lagrange equations.*

Euler-Poincaré reduction. Let us begin with the important, but special cases of Euler-Poincaré reduction and Lie-Poisson reduction. Following this we will comment on generalizations. Let G be a Lie group (finite dimensional for simplicity) and let G act on itself by left translation and hence, by tangent lift, on its tangent bundle TG . Let $L : TG \rightarrow \mathbb{R}$ be a G -invariant Lagrangian. Being invariant, L is completely determined by its restriction to

the tangent space at the identity e . Identifying $T_e G$, the tangent space to G at e with the Lie algebra \mathfrak{g} of G , we let $l : \mathfrak{g} \rightarrow \mathbb{R}$ be defined by $l = L|_{T_e G}$. Alternatively, we may identify the quotient space TG/G with \mathfrak{g} and l is the map induced by L on the quotient space.

The *velocity* of the system is given by $\dot{g}(t)$, thought of as a tangent vector to G at $g(t)$. The *body velocity* is defined by $\xi(t) = g(t)^{-1}\dot{g}(t)$, the left translation of \dot{g} to the identity.

Theorem 1.1 (Euler-Poincaré reduction-part 1.) *A curve $g(t)$ in G satisfies the Euler-Lagrange equations for L iff $\xi(t)$ satisfies the **Euler-Poincaré equations** for l :*

$$\frac{d}{dt} \frac{\partial l}{\partial \xi} = \text{ad}_\xi^* \frac{\partial l}{\partial \xi}.$$

Let me explain the notation: $\partial l / \partial \xi$ means the differential of l , so it is an element of \mathfrak{g}^* (and it is understood that $\partial l / \partial \xi$ is evaluated at the point $\xi(t)$). The **ad map** is defined by Lie algebra bracketing:

$$\text{ad}_\xi : \mathfrak{g} \rightarrow \mathfrak{g}; \quad \eta \mapsto [\xi, \eta]$$

and $\text{ad}_\xi^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ means its dual map.

There are many ways to prove this theorem, but the technique of *reduction of variational principles* is very efficient as well as carrying an important message. It proceeds by asking if we can write Hamilton's principle entirely in terms of l and ξ . By left invariance,

$$L(g, \dot{g}) = l(\xi),$$

so this much is easy. Variations are a little trickier but to help matters, we shall assume G is a matrix group. Take variations of the *reconstruction equation* (so called because it allows one to determine g from ξ):

$$\xi = g^{-1}\dot{g}$$

to give

$$\delta \xi = -g^{-1}\delta g g^{-1}\dot{g} + g^{-1}(\delta g) \dot{}.$$

As in the calculus of variations, a variation such as $\delta \xi$ or δg is nothing more than the derivative of a parameterized family of curves with fixed endpoints, with respect to this parameter. Let

$$\zeta = g^{-1}\delta g$$

i.e., ζ is the variation of g left translated to the identity. We refer to ζ as the **body representation** of the variation.

Since the time derivative of g^{-1} is $-g^{-1}\dot{g}g^{-1}$, we get

$$\dot{\zeta} = -g^{-1}\dot{g}g^{-1}\delta g + g^{-1}(\delta g) \dot{},$$

so that

$$\delta \xi = -\zeta \xi + \xi \zeta + \dot{\zeta} = \dot{\zeta} + [\xi, \zeta].$$

Thus, variations of ξ induced by variations of g must have this form for some curve $\zeta \in \mathfrak{g}$ vanishing at the endpoints. This calculation shows that

Theorem 1.2 (Euler-Poincaré reduction-part 2.) *Hamilton's principle for L on G is equivalent to the **reduced variational principle**:*

$$\delta \mathfrak{S}_{\text{red}} = \delta \int l(\xi(t)) dt = 0$$

for variations of the form

$$\delta \xi = \dot{\zeta} + [\xi, \zeta]$$

for some curve ζ in \mathfrak{g} vanishing at the endpoints.

Here, $\mathfrak{S}_{\text{red}}$ is the **reduced action**, the integral of the reduced Lagrangian l along curves in \mathfrak{g} . Now it is easy to work out the corresponding equations as one does in the calculus of variations and this yields a proof of part 1 of the Euler-Poincaré reduction.

History and literature. Looking back with modern notation, it is fair to say that Lagrange [1788] realized the importance of relating the dynamics on TG to the dynamics on \mathfrak{g} and devoted much of Volume II of *Mécanique Analytique* to it for the case of the rotation group $\text{SO}(3)$. The Euler-Poincaré equations for *general Lie algebras* were first written down by Poincaré [1901b] who realized that they were fundamental equations in fluid and solid mechanics, as is apparent from Poincaré [1910]. However, it seems that Poincaré did not take a variational point of view. Arnold [1966a] developed the geometry and mechanics of the Euler-Poincaré equations, including stability theory. In fluid mechanics, the constraints on the variations appearing in reduced variational principles go under the name “Lin constraints” and this subject has a large and complex literature. The general theorem presented here is due to Marsden and Scheurle [1993b] and, for general Lie groups, to Bloch, Krishnaprasad, Marsden and Ratiu [1996]. The Euler-Poincaré equations in a general context including advected parameters are developed in Holm, Marsden and Ratiu [1998a], linked to semidirect product reduction theory and applied to continuum mechanics.

Example 1. The first example is the rigid body, free to spin about its center of mass. Here we take $G = \text{SO}(3)$ so that $\mathfrak{g} \cong \mathbb{R}^3$ with Lie algebra bracket given by the cross product. The Lagrangian L is the total kinetic energy. The reduced Lagrangian $l : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a quadratic function of ξ , which in this case is called the **body angular velocity** and is written Ω . Thus,

$$l(\Omega) = \frac{1}{2} \langle \Omega, I\Omega \rangle$$

for a symmetric positive definite matrix I , the **moment of inertia tensor**. The Euler-Poincaré equations (or the **rigid body equations**) are readily seen to be

$$I\dot{\Omega} = (I\Omega) \times \Omega.$$

The reader may, as an exercise, check directly that these equations come from the reduced variational principle. Since l is quadratic, L is the kinetic energy of a Riemannian metric on $\text{SO}(3)$, so the Euler-Lagrange equations on $\text{SO}(3)$ are the geodesic equations.

Example 2. For the motion of an ideal fluid, we choose $G = \text{Diff}_{\text{vol}}(\mathcal{B})$, the group of volume preserving diffeomorphisms of a given domain \mathcal{B} , a Riemannian manifold in which the fluid moves. The Lagrangian L is the total fluid kinetic energy. However, this time L is

right invariant and the reduced Lagrangian is defined on the Lie algebra \mathfrak{g} , the divergence free vector fields on \mathcal{B} , by

$$l(u) = \frac{1}{2} \int \|u\|^2 d\mu,$$

where $\|\cdot\|^2$ denotes the Riemannian norm on \mathcal{B} and $d\mu$ is the associated volume form. Note again l is quadratic, so we get geodesics. The Euler-Poincaré equations are the *ideal fluid equations*

$$\frac{\partial u}{\partial t} + \nabla_u u = -\nabla p$$

where the pressure p is determined implicitly by the constraint $\operatorname{div} u = 0$. Again the reader should work out the reduced variational principle and verify that these equations are of Euler-Poincaré form. Relevant references are Arnold [1966a], Ebin and Marsden [1970], Marsden and Ratiu [1998].

Example 3. The dynamics of a rigid body in a fluid are often modeled by the classical Kirchhoff equations in which the fluid is assumed to be potential flow, responding to the motion of the body. (For underwater vehicle dynamics we will need to include buoyancy effects.)¹ Here we choose $G = \operatorname{SE}(3)$, the group of Euclidean motions of \mathbb{R}^3 and the Lagrangian is the total energy of the body-fluid system. Recall that the Lie algebra of $\operatorname{SO}(3)$ is $\mathfrak{se}(3) = \mathbb{R}^3 \times \mathbb{R}^3$ with the bracket

$$[(\Omega, u), (\Sigma, v)] = (\Omega \times \Sigma, \Omega \times v - \Sigma \times u).$$

The reduced Lagrangian is again quadratic, so has the form

$$l(\Omega, v) = \frac{1}{2} \Omega^T J \Omega + \Omega^T Dv + \frac{1}{2} v^T Mv.$$

The Euler-Poincaré equations are computed to be

$$\left. \begin{aligned} \dot{\Pi} &= \Pi \times \Omega + P \times v \\ \dot{P} &= P \times \Omega \end{aligned} \right\}$$

where $\Pi = \partial l / \partial \Omega = J\Omega + Dv$ the “angular momentum” and $P = \partial l / \partial v = Mv + D^T \Omega$, the “linear momentum”.

Again, we suggest that the reader work out the reduced variational principle. Relevant references are Lamb [1932], Leonard [1996], Leonard and Marsden [1997], and Holmes, Jenkins and Leonard [1997].

Example 4. Following Ovsienco and Khesin [1987], we will now indicate how the KdV equations may be recast as Euler-Poincaré equations. The *KdV equation* is the following equation for a scalar function $u(x, t)$ of the real variables x and t :

$$u_t + 6uu_x + u_{xxx} = 0.$$

We let \mathfrak{g} be the Lie algebra of vector fields u on the circle (of length 1) with the standard bracket

$$[u, v] = u'v - v'u.$$

¹This model may be viewed inside the larger model of an elastic-fluid interacting system with the constraint of rigidity imposed on the elastic body and with the reduced space for the fluid variables (potential flow is simply reduction at zero for fluids).

Let the *Gelfand-Fuchs cocycle* be defined by²

$$\Sigma(u, v) = \gamma \int_0^1 u'(x)v''(x)dx,$$

where γ is a constant. Let the *Virasoro Lie algebra* be defined by $\tilde{\mathfrak{g}} = \mathfrak{g} \times \mathbb{R}$ with the Lie bracket

$$[(u, a), (v, b)] = ([u, v], \gamma\Sigma(u, v)).$$

This is verified to be a Lie algebra; the corresponding group is called the *Bott-Virasoro group*. Let

$$l(u, a) = \frac{1}{2}a^2 + \int_0^1 u^2(x)dx.$$

Then one checks that the Euler-Poincaré equations are

$$\begin{aligned} \frac{da}{dt} &= 0 \\ \frac{du}{dt} &= -\gamma au''' - 3u'u \end{aligned}$$

so that for appropriate a and γ and rescaling, we get the KdV equation. Thus, *the KdV equations may be regarded as geodesics on the Bott-Virasoro group*.

Likewise, the Camassa-Holm equation can be recast as geodesics using the H^1 rather than the L^2 metric (see Misiolek [1997] and Holm, Kouranbaeva, Marsden, Ratiu and Shkoller [1998]).

Lie-Poisson Systems. The Euler-Poincaré equations occur on \mathfrak{g} , while the Lie-Poisson equations occur on \mathfrak{g}^* . To understand how one arrives at \mathfrak{g}^* we recall a few facts about reduction.

If P is a Poisson manifold and G acts on it freely and properly, then P/G is also Poisson in a natural way: identify functions on P/G with G -invariant functions on P and use this to induce a bracket on functions on P/G . In the case $P = T^*G$ and G acts on the left by cotangent lift, then $T^*G/G \cong \mathfrak{g}^*$ inherits a Poisson structure given explicitly by the following theorem.

Theorem 1.3 (Lie-Poisson reduction.) *The Poisson structure inherited on \mathfrak{g}^* is given by*

$$\{f, g\}_-(\mu) = - \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu} \right] \right\rangle,$$

the Lie-Poisson bracket. For the right action, use $+$.

*If H is G -invariant on T^*G and X_H is its Hamiltonian vector field determined by $\dot{F} = \{F, H\}$, then X_H projects to the Hamiltonian vector field X_h determined by $\dot{f} = \{f, h\}_-$ where $h = H|_{T_e^*G} = H|_{\mathfrak{g}^*}$. We call $\dot{f} = \{f, h\}_-$ the **Lie-Poisson equations**.*

If l is regular; *i.e.*, $\xi \mapsto \mu = \partial l / \partial \xi$ is invertible, then the Legendre transformation taking ξ to μ and l to

$$h(\mu) = \langle \xi, \mu \rangle - l(\xi)$$

²An interesting interpretation of the Gelfand-Fuchs cocycle as the curvature of a mechanical connection is given in Marsden, Misiolek, Perlmutter and Ratiu [1998a,b].

maps the Euler-Poincaré equations to the Lie-Poisson equations and vice-versa.³

Lie-Poisson systems have a remarkable property; they leave the coadjoint orbits in \mathfrak{g}^* invariant. In fact the coadjoint orbits are the symplectic leaves of \mathfrak{g}^* . For each of examples 1 and 3, the reader may check directly that the equations are Lie-Poisson and that the coadjoint orbits are preserved. For example 2, the preservation of coadjoint orbits is essentially Kelvin's circulation theorem. See Marsden and Weinstein [1983] for details. For example 1, the coadjoint orbits are the familiar momentum spheres, shown in figure 1.1.

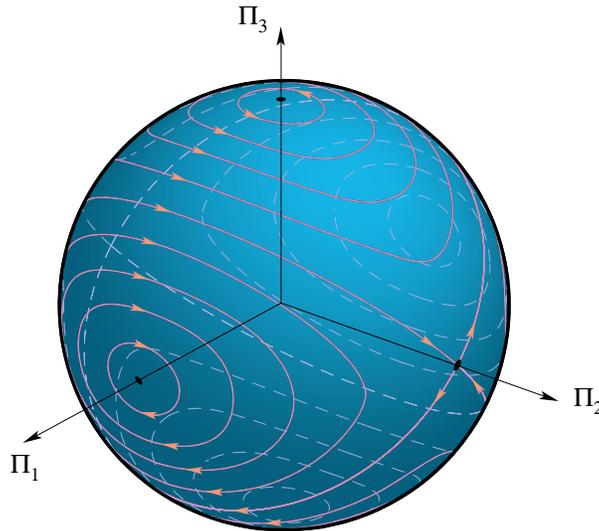


Figure 1.1: The rigid body momentum sphere.

History and literature. Lie-Poisson brackets were known to Lie around 1890, but apparently this aspect of the theory was not picked up by Poincaré. The coadjoint orbit symplectic structure was discovered by Kirillov, Kostant and Souriau in the 1960's. They were shown to be symplectic reduced spaces by Marsden and Weinstein [1974]. It is not clear who first observed *explicitly* that \mathfrak{g}^* inherits the Lie-Poisson structure by reduction as in the preceding Lie-Poisson reduction theorem. It is *implicit* in many works such as Lie [1890], Kirillov [1962], Guillemin and Sternberg [1980] and Marsden and Weinstein [1982, 1983], but is *explicit* in Holmes and Marsden [1983] and Marsden, Weinstein, Ratiu, Schmid and Spencer [1983].

Symplectic and Poisson reduction. The ways in which reduction has been generalized and applied has been nothing short of phenomenal. Let me sketch just a few of the highlights (eliminating many important references).

First of all, in an effort to synthesize coadjoint orbit reduction (suggested by work of Arnold, Kirillov, Kostant and Souriau) with techniques for the reduction of cotangent bundles by Abelian groups of Smale [1970], Marsden and Weinstein [1974] developed symplectic reduction; related results, but with a different motivation and construction were found by

³A cautionary note. The heavy top is an example of a Lie-Poisson system on $\mathfrak{se}(3)^*$. However, its inverse Legendre transformation (using the standard h) is degenerate! This is an indication that something is missing on the Lagrangian side and this is indeed the case. The resolution is found in Holm, Marsden and Ratiu [1998a].

Meyer [1973]. The construction is now well known: let (P, Ω) be a symplectic manifold and $J : P \rightarrow \mathfrak{g}^*$ be an equivariant momentum map; then avoiding singularities, $J^{-1}(\mu)/G_\mu = P_\mu$ is a symplectic manifold in a natural way. For example, for $P = T^*G$, one gets coadjoint orbits.

Kazhdan, Kostant and Sternberg [1978] showed how P_μ can be realized in terms of orbit reduction $P_\mu \cong J^{-1}(\mathcal{O})/G$ and from this it follows that P_μ are the symplectic leaves in P/G . This paper was also one of the first to notice deep links between reduction and integrable systems, a subject continued by, for example, Bobenko, Reyman and Semenov-Tian-Shansky [1989].

The way in which the *Poisson* structure on P_μ is related to that on P/G was clarified in a generalization of Poisson reduction due to Marsden and Ratiu [1986], a technique that has also proven useful in integrable systems (see, for example, Pedroni [1995] and Vanhaecke [1996]).

The mechanical connection. A basic construction implicit in Smale [1970], Abraham and Marsden [1978] and explicit in Kummer [1981] is the notion of the mechanical connection. The geometry of this situation was used to great effect in Guichardet [1984] and Iwai [1987, 1990].

Assume Q is Riemannian (the metric often being the kinetic energy metric) and that G acts on Q freely by isometries, so $\pi : Q \rightarrow Q/G$ is a principal bundle. If we declare the horizontal spaces to be metric orthogonal to the group orbits, this uniquely defines a connection called the *mechanical connection*. There are explicit formulas for it in terms of the locked inertia tensor; see for instance, Marsden [1992] for details. The space Q/G is called *shape space* and plays a critical role in the theory.⁴

Tangent and cotangent bundle reduction. The simplest case of cotangent bundle reduction is reduction at zero in which case one has $(T^*Q)_{\mu=0} = T^*(Q/G)$, the latter with the canonical symplectic form. Another basic case is when G is abelian. Here, $(T^*Q)_\mu \cong T^*(Q/G)$ but the latter has a symplectic structure modified by magnetic terms; that is, by the curvature of the mechanical connection.

The Abelian version of cotangent bundle reduction was developed by Smale [1970] and Sater [1975] and was generalized to the nonabelian case in Abraham and Marsden [1978]. It was Kummer [1981] who introduced the interpretations of these results in terms of the mechanical connection.

The Lagrangian analogue of cotangent bundle reduction is called *Routh reduction* and was developed by Marsden and Scheurle [1993a,b]. Routh, around 1860 investigated what we would call today the Abelian version.

The “bundle picture” begun by the developments of the cotangent bundle reduction theory was significantly developed by Montgomery, Marsden and Ratiu [1984] and Montgomery [1986] motivated by work of Weinstein and Sternberg on Wong’s equations (the equations for a particle moving in a Yang-Mills field).

This bundle picture can be viewed as follows. Choosing a connection, such as the mechanical connection, on $Q \rightarrow Q/G$, one gets a natural isomorphism

$$T^*Q/G \cong T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$$

where the sum is a Whitney sum of vector bundles over Q/G (fiberwise a direct sum) and $\tilde{\mathfrak{g}}^*$ is the associated vector bundle to the co-adjoint action of G on \mathfrak{g}^* . The description of the Poisson structure on this bundle (a synthesis of the canonical bracket, the Lie-Poisson bracket and curvature) may be found in Cendra, Marsden and Ratiu [1998].

⁴Shape space and its geometry plays a key role in computer vision. See for example, Le and Kendall [1993].

Lagrangian reduction. The Lagrangian analogue of the bundle picture is the dual isomorphism

$$TQ/G \cong T(Q/G) \oplus \tilde{\mathfrak{g}}$$

whose geometry is developed in Cendra, Marsden and Ratiu [1998]. In particular, the equations and variational principles are developed on this space. For $Q = G$ this reduces to the Euler-Poincaré picture we had previously. For G abelian, it reduces to the Routh procedure.

If we have an invariant Lagrangian on TQ it induces a Lagrangian l on $(TQ)/G$ and hence on $T(Q/G) \oplus \tilde{\mathfrak{g}}$. Calling the variables r^α, \dot{r}^α and Ω^α , the resulting **reduced Euler-Lagrange** equations (implicitly contained in Cendra, Ibort and Marsden [1987] and explicitly in Marsden and Scheurle [1993b]) are

$$\begin{aligned} \frac{d}{dt} \frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} &= \frac{\partial l}{\partial \Omega^\alpha} (-B_{\alpha\beta}^\alpha \dot{r}^\beta + \xi_{\alpha d}^a \Omega^d) \\ \frac{d}{dt} \frac{\partial l}{\partial \Omega^b} &= \frac{\partial l}{\partial \Omega^a} (-\xi_{\alpha\beta}^a \dot{r}^\alpha + C_{ab}^a \Omega^d) \end{aligned}$$

where $B_{\alpha\beta}^a$ is the curvature of the connection \mathcal{A}_α^b , C_{bd}^a are the structure constants of the Lie algebra \mathfrak{g} and where $\xi_{\alpha d}^a = C_{bd}^a \mathcal{A}_\alpha^b$.

Using the geometry of the bundle $TQ/G = T(Q/G) \oplus \tilde{\mathfrak{g}}$, one obtains a nice interpretation of these equations in terms of covariant derivatives. One easily gets the dynamics of particles in a Yang-Mills field (these are called Wong's equation) as a special case; see Cendra, Holm, Marsden and Ratiu [1998] for this example. Wong's equations will come up again in Lecture 4 on optimal control.

We also mention that methods of Lagrangian reduction have proven very useful in optimal control problems. It was used in Koon and Marsden [1997] to extend the falling cat theorem of Montgomery [1990] to the case of nonholonomic systems.

Cotangent bundle reduction is very interesting for group extensions, such as the Bott-Virasoro group described earlier, where the Gelfand-Fuchs cocycle may be interpreted as the curvature of a mechanical connection. This is closely related to work of Marsden, Misiolek, Perlmutter and Ratiu [1998a,b] on reduction by stages. This work in turn is an outgrowth of earlier work of Guillemin and Sternberg [1980], Marsden, Ratiu and Weinstein [1984a,b] and many others on systems such as the heavy top, compressible flow and MHD. It also applies to underwater vehicle dynamics as shown in Leonard [1996] and Leonard and Marsden [1997].

In semidirect product reduction, one supposes that G acts on a vector space V (and hence on its dual V^*). From G and V we form the **semidirect product** Lie group $S = G \ltimes V$, the set $G \times V$ with multiplication

$$(g_1, v_1) \cdot (g_2, v_2) = (g_1 g_2, v_1 + g_1 v_2).$$

The Euclidean group $SE(3) = SO(3) \ltimes \mathbb{R}^3$, the semidirect product of rotations and translations is a basic example. Now suppose we have a Hamiltonian on T^*G that is invariant under the isotropy G_{a_0} for $a_0 \in V^*$. The **semidirect product reduction theorem** states that *reduction of T^*G by G_{a_0} gives reduced spaces that are symplectically diffeomorphic to coadjoint orbits in the dual of the Lie algebra of the semi-direct product: $(\mathfrak{g} \ltimes V)^*$.*

This is a very important construction in applications where one has "advected quantities" (such as density in compressible flow). Its Lagrangian counterpart, which is *not* simply the Euler-Poincaré equations on $\mathfrak{g} \ltimes V$, is developed in Holm, Marsden and Ratiu [1998a] along with applications to continuum mechanics. Cendra, Holm, Hoyle and Marsden [1998] have applied this idea to the Maxwell-Vlasov equations of plasma physics.

If one reduces the semidirect product group $S = G \ltimes V$ in two stages, first by V and then by G , one recovers the semidirect product reduction theorem mentioned above.

A far reaching generalization of this semidirect product theory is given in Marsden, Misiolek, Perlmutter and Ratiu [1998a,b] in which one has a group M with a normal subgroup $N \subset M$ and M acts on a symplectic manifold P . One wants to reduce P in two stages, first by N and then by M/N . On the Poisson level this is easy: $P/M \cong (P/N)/(M/N)$ but on the symplectic level it is quite subtle. Cendra, Marsden and Ratiu [1998] have developed a Lagrangian counterpart to reduction by stages.

Singular reduction. Singular reduction starts with the observation of Smale [1970] that $z \in P$ is a regular point of \mathbf{J} iff z has no continuous isotropy. Motivated by this, Arms, Marsden and Moncrief [1981] showed that the level sets $\mathbf{J}^{-1}(0)$ of an equivariant momentum map \mathbf{J} have quadratic singularities at points with continuous symmetry. While easy for compact group actions, their main examples were infinite dimensional! The structure of $\mathbf{J}^{-1}(0)/G$ for compact groups was developed in Sjamaar and Lerman [1990], and extended to $\mathbf{J}^{-1}(\mu)/G_\mu$ by Bates and Lerman [1996] and Ortega and Ratiu [1997a]. Many specific examples of singular reduction and further references may be found in Bates and Cushman [1997].

The method of invariants. An important method for the reduction construction is called the *method of invariants*. This method seeks to parameterize quotient spaces by functions that are invariant under the group action. The method has a rich history going back to Hilbert's *invariant theory* and it has much deep mathematics associated with it. It has been of great use in bifurcation with symmetry (see Golubitsky, Stewart and Schaeffer [1988] for instance).

In mechanics, the method was developed by Kummer, Cushman, Rod and coworkers in the 1980's. We will not attempt to give a literature survey here, other than to refer to Kummer [1990], Kirk, Marsden and Silber [1996] and the book of Bates and Cushman [1997] for more details and references. We shall illustrate the method with a famous system, the three wave interaction, based on Alber, Luther, Marsden and Robbins [1998b].

The three wave interaction. The *quadratic resonant three wave equations* are the following ode's on \mathbb{C}^3 :

$$\begin{aligned}\frac{dq_1}{dt} &= i s_1 \gamma_1 q_2 \bar{q}_3 \\ \frac{dq_2}{dt} &= i s_2 \gamma_2 q_1 q_3 \\ \frac{dq_3}{dt} &= i s_3 \gamma_3 \bar{q}_1 q_2\end{aligned}$$

Here, $q_1, q_2, q_3 \in \mathbb{C}$, $i = \sqrt{-1}$, the overbar means complex conjugate, and γ_1, γ_2 and γ_3 are nonzero real numbers with $\gamma_1 + \gamma_2 + \gamma_3 = 0$. The choice $(s_1, s_2, s_3) = (1, 1, -1)$; gives the *decay interaction*, while $(s_1, s_2, s_3) = (-1, 1, 1)$ gives the *explosive interaction*.

Resonant wave interactions describe energy exchange among nonlinear modes in contexts involving nonlinear waves (the Benjamin-Feir instability, etc.) in fluid mechanics, plasma physics and other areas. There are other versions of the equations in which coupling associated with phase modulations appears through linear and cubic terms. Much of our motivation comes from nonlinear optics (optical transmission and switching). The three wave equations are discussed in, for example, Whitham [1974] and its dynamical systems aspects are explored in Guckenheimer and Mahalov [1992].

The methods we develop work rather generally for resonances—the rigid body is well known to be intimately connected with the 1:1 resonance (see, for example, Cushman and Rod [1982], Churchill, Kummer and Rod [1983]). The three wave interaction has an interesting Hamiltonian and integrable structure. We shall use a standard Hamiltonian structure and the technique of invariants to understand it. The decay system is *Lie-Poisson* for the Lie algebra $\mathfrak{su}(3)$ – this is the one of notable interest for phases (the explosive case is associated with $\mathfrak{su}(2, 1)$). This is related to the Lax representation of the equations—the n -wave interaction is likewise related to $\mathfrak{su}(n)$. The general picture developed is useful for many other purposes, such as polarization control (building on work of David, Holm and Tratnik [1989] and David and Holm [1990]) and perturbations of Hamiltonian normal forms (see Kirk, Marsden and Silber [1996]).

The canonical Hamiltonian structure. We describe how the three wave system is Hamiltonian relative to a canonical Poisson bracket. We choose (primarily a matter of convenience) a γ_i -weighted canonical bracket on \mathbb{C}^3 . This bracket has the real and imaginary parts of each complex dynamical variable q_i as conjugate variables. Correspondingly, we will use a cubic Hamiltonian. The *scaled canonical Poisson bracket* on \mathbb{C}^3 may be written in complex notation as

$$\{F, G\} = -2i \sum_{k=1}^3 s_k \gamma_k \left(\frac{\partial F}{\partial q^k} \frac{\partial G}{\partial \bar{q}_k} - \frac{\partial G}{\partial q^k} \frac{\partial F}{\partial \bar{q}_k} \right).$$

The corresponding *symplectic structure* can be written

$$\Omega((z_1, z_2, z_3), (w_1, w_2, w_3)) = - \sum_{k=1}^3 \frac{1}{s_k \gamma_k} \text{Im}(z_k \bar{w}_k).$$

The (cubic) *Hamiltonian* is

$$H = -\frac{1}{2} (\bar{q}_1 q_2 \bar{q}_3 + q_1 \bar{q}_2 q_3).$$

Hamilton's equations for a Hamiltonian H are

$$\frac{dq_k}{dt} = \{q_k, H\},$$

and it is straightforward to check that Hamilton's equations are given in complex notation by

$$\frac{dq_k}{dt} = -2i s_k \gamma_k \frac{\partial H}{\partial \bar{q}_k},$$

One checks that Hamilton's equations in our case coincide with the three wave equations.

Integrals of motion. Besides H itself, there are additional constants of motion, often referred to as the *Manley-Rowe relations*:

$$\begin{aligned} K_1 &= \frac{|q_1|^2}{s_1 \gamma_1} + \frac{|q_2|^2}{s_2 \gamma_2}, \\ K_2 &= \frac{|q_2|^2}{s_2 \gamma_2} + \frac{|q_3|^2}{s_3 \gamma_3}, \\ K_3 &= \frac{|q_1|^2}{s_1 \gamma_1} - \frac{|q_3|^2}{s_3 \gamma_3}. \end{aligned}$$

The vector function (K_1, K_2, K_3) is the *momentum map* for the following symplectic action of the group $T^3 = S^1 \times S^1 \times S^1$ on \mathbb{C}^3 :

$$\begin{aligned}(q_1, q_2, q_3) &\mapsto (q_1 \exp(i\gamma_1), q_2 \exp(i\gamma_1), q_3), \\(q_1, q_2, q_3) &\mapsto (q_1, q_2 \exp(i\gamma_2), q_3 \exp(i\gamma_2)), \\(q_1, q_2, q_3) &\mapsto (q_1 \exp(i\gamma_3), q_2, q_3 \exp(-i\gamma_3)).\end{aligned}$$

The Hamiltonian taken with any two of the K_j are checked to be a complete and independent set of conserved quantities. Thus, the system is *Liouville-Arnold integrable*.

The K_j clearly give only two independent invariants since $K_1 - K_2 = K_3$. Any combination of two of these actions can be generated by the third reflecting the fact that the K_j are linearly dependent. Another way of saying this is that the group action by T^3 is really captured by the action of T^2 .

Integrating the equations. To carry out the integration, one can make use of the Hamiltonian plus two of the integrals, K_j to reduce the system to quadratures. This is often carried out using the transformation $q_j = \sqrt{\rho_j} \exp i\phi_j$ to obtain expressions for the phases ϕ_j . The resulting expressions are nice, but the alternative point of view using invariants is also useful.

Poisson reduction. Symplectic reduction of the above Hamiltonian system uses the symmetries and associated conserved quantities K_k . In *Poisson reduction*, we replace \mathbb{C}^3 with the *orbit space* \mathbb{C}^3/T^2 , which then inherits a Poisson structure. To obtain the *symplectic leaves* in this reduction, we use the *method of invariants*. Invariants for the T^2 action are:

$$\begin{aligned}X + iY &= q_1 \bar{q}_2 q_3 \\Z_1 &= |q_1|^2 - |q_2|^2 \\Z_2 &= |q_2|^2 - |q_3|^2\end{aligned}$$

These quantities provide coordinates for the four dimensional orbit space \mathbb{C}^3/T^2 . The following identity (this is part of the invariant theory game) holds for these invariants and the conserved quantities:

$$X^2 + Y^2 = \beta(\delta - Z_2)(Z_2 + s_3\gamma_3 K_2)(s_2\gamma_2 K_2 - Z_2)$$

where the constants β, δ are given by

$$\beta = \frac{s_1\gamma_1 s_2\gamma_2 s_3\gamma_3}{(s_2\gamma_2 + s_3\gamma_3)^3}, \quad \delta = s_2\gamma_2 K_1 + s_3\gamma_3 (K_1 - K_2).$$

This defines a two dimensional surface in (X, Y, Z_2) space, with Z_1 determined by the values of these invariants and the conserved quantities (so it may also be thought of as a surface in (X, Y, Z_1, Z_2) as well). A sample of one of these surfaces is plotted in Figure 1.2.

We call these surfaces the *three wave surfaces*. They are examples of *orbifolds*. The evident singularity in the space is typical of orbifolds and comes about from the non-freeness of the group action.

Any trajectory of the original equations defines a curve on each three wave surface, in which the K_j are set to constants. These three wave surfaces are the symplectic leaves in the four dimensional Poisson space with coordinates (X, Y, Z_1, Z_2) .

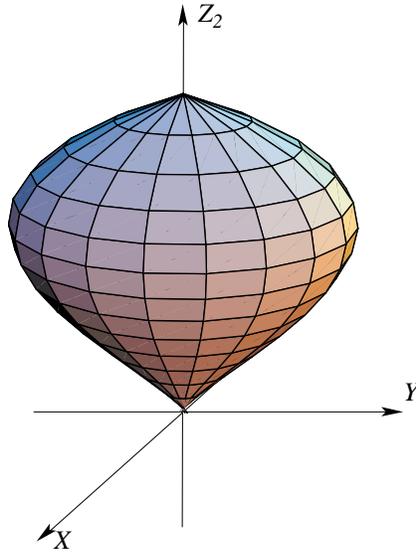


Figure 1.2: The reduced phase space for the three-wave equations.

The original equations define a dynamical system in the Poisson reduced space and on the symplectic leaves as well. The *reduced Hamiltonian* is

$$H(X, Y, Z_1, Z_2) = -X$$

and indeed, $\dot{X} = 0$ is one of the reduced equations. Thus, the trajectories on the reduced surfaces are obtained by slicing the surface with the planes $X = \text{Constant}$. The *Poisson structure* on \mathbb{C}^3 drops to a Poisson structure on (X, Y, Z_1, Z_2) -space, and the *symplectic structure* drops to one on each three wave surface—this is of course an example of the general procedure of *symplectic reduction*. Also, from the geometry, it is clear that interesting *homoclinic orbits* pass through the singular points—these are cut out by the plane $X = 0$.

A *control perspective* allows one to manipulate the plane $H = -X$ and thereby the dynamics. This aspect is explored in Alber, Luther, Marsden and Robbins [1998a].

Lecture 2

Stability, Underwater Vehicle Dynamics and Phases

Some history, background and literature on stability. The *energy momentum method* is an extension of the Arnold (or energy-Casimir) method for the study of stability of relative equilibria for Lie-Poisson systems on duals of Lie algebras, especially those of fluid dynamical type. The method simultaneously extends and refines the fundamental stability techniques going back to Routh, Liapunov and in more recent times, to Smale [1970].

There are several motivations for developing these extensions. First of all, the energy-momentum method can deal with Lie-Poisson systems for which there are not sufficient Casimir functions available for the Arnold method to be effective, such as 3D ideal flow and certain problems in elasticity. We recall that in the Arnold method, one seeks a Casimir function C (plus possibly other conserved quantities) such that $H + C$ has a critical point at the equilibrium in question and that $\delta^2(H + C)$ is definite there (or satisfies suitable convexity assumptions). For 3D Euler flow there is only one known Casimir function, the helicity, which is not enough to analyze the stability of most equilibria—even at the first variation step. Abarbanel and Holm [1987] use what we see retrospectively is the energy-momentum method to show that 3d equilibria for ideal flow are always formally unstable due to vortex stretching. Other fluid and plasma situations, such as those considered by Chern and Marsden [1990] for ABC flows, and certain multiple hump situations in plasma dynamics (see Holm, Marsden, Ratiu and Weinstein [1985] and Morrison [1987] for example) provided additional motivation for generalizations in the Lie-Poisson setting.

A second motivation is to extend the method to systems that need not be Lie-Poisson and still make use of the powerful ideas of reduction (even if reduced spaces are not *literally* used), as in the original Arnold method. Examples such as rigid bodies with vibrating antennas (Sreenath, et al [1988], Oh et al [1989], Krishnaprasad and Marsden [1987]) and coupled rigid bodies (Patrick [1989]) motivated the need for such an extension of the theory.

Finally, it gives *sharper stability conclusions* in material representation and links with geometric phases, as we shall discuss.

The idea of the energy-momentum method. The setting of the energy-momentum method is that of phase space P with a symmetry group G , a G -invariant Hamiltonian H and an associated conserved momentum \mathbf{J} , an equivariant momentum map, $\mathbf{J} : P \rightarrow \mathfrak{g}^*$.

A *relative equilibrium* (also called a *steady motion*) is a point $z_e \in P$ whose dynamical orbit $z(t)$ is coincident with a one-parameter group orbit: $z(t) = \exp(t\xi)z_e$ for some $\xi \in \mathfrak{g}$.

Assumptions. The standard energy-momentum method assumes that the relative equilibrium z_e is a *regular point* for \mathbf{J} and that the image point $\mu_e = \mathbf{J}(z_e)$ is a *generic, i.e., nonspecial* point. Patrick [1992], extended this result, allowing μ_e to be nongeneric. A crucial hypothesis in Patrick's result is a compactness condition on the isotropy subgroup G_{μ_e} , the subgroup of elements of G that fix μ_e . The result is that one gets stability modulo G_{μ_e} .

The rough idea for the energy momentum method is to first formulate the problem directly on the unreduced space. Here, relative equilibria associated with a Lie algebra element ξ are critical points of the augmented Hamiltonian $H_\xi := H - \langle \mathbf{J}, \xi \rangle$. Next, compute the second variation of H_ξ at a relative equilibria z_e with momentum value μ_e subject to the constraint $J = \mu_e$ and on a space, say E_{z_e} , transverse to the action of G_{μ_e} (see Figure 2.1).

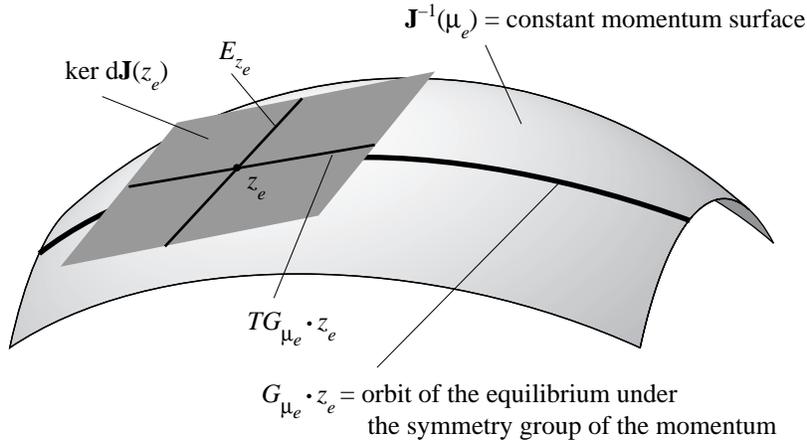


Figure 2.1: The geometry of the energy-momentum method.

This, as a *general method*, was first given in Marsden Simo, Lewis and Posburgh [1989] and Marsden and Simo [1990]. To extend the method to the singular case one replaces the above figure by the corresponding *symplectic slice* (see Arms, Marsden and Moncrief [1981], Guillemin and Sternberg [1984] and Ortega and Ratiu [1997b]).

Although the augmented Hamiltonian plays the role of $H + C$ in the Arnold method, notice that Casimir functions are *not required* to carry out the calculations.

Block diagonalization. The theory for carrying out this procedure was much developed in Simo, Posbergh and Marsden [1990, 1991] and Simo, Lewis and Marsden [1991] as we shall briefly explain. An exposition of the method may be found, along with additional references in Marsden [1992]. To obtain the more detailed structures found in these papers, we consider the case when $P = T^*Q$ for a configuration space Q a symmetry group G acting on Q , with the standard cotangent bundle momentum map $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$, where \mathfrak{g}^* is the Lie algebra of G . Thus, we have the well known formula $\langle \mathbf{J}(\alpha_q), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle$. Of course one gets the Lie-Poisson case when $Q = G$. In this cotangent bundle case and when H is kinetic minus potential, an *amazing thing happens*: using splittings that are based on the mechanical connection, the second variation of H_ξ at the relative equilibrium can always be arranged to be block diagonal, while, *simultaneously*, the symplectic structure also has a simple block structure so that the linearized equations are put into a useful canonical form. Even in the Lie-Poisson setting, this leads to situations in which one gets

much simpler second variations. *This block diagonal structure is what gives the method its computational power.* Roughly speaking, for rotating systems, this method optimally separates the rotational and vibrational modes. In fact, links between these methods and problems in molecular dynamics are most interesting (see Littlejohn and Reich [1996]).

Lagrangian version of the energy-momentum method. The energy momentum method may also be usefully formulated in the Lagrangian setting and this setting is very convenient for the calculations in many examples. The general theory for this was done in Lewis [1992b] and Wang and Krishnaprasad [1992]. This Lagrangian setting is closely related to the general theory of Lagrangian reduction mentioned in Lecture 1. In this context one reduces variational principles rather than symplectic and Poisson structures and for the case of reducing the tangent bundle of a Lie group, recall that this leads to the Euler-Poincaré equations.

Effectiveness in examples. The energy momentum method has proven its effectiveness in a number of examples. For instance, Lewis and Simo [1990] were able to deal with the stability problem for pseudo-rigid bodies, which was thought up to that time to be analytically intractable.

The energy-momentum method can sometimes be used in contexts where the reduced space is singular or at nongeneric points in the dual of the Lie algebra. This is done at singular points in Lewis [1992b] and Lewis, Ratiu, Simo and Marsden [1992] who analyze the heavy top. As we mentioned above, it was extended for compact groups, to allow nongeneric points $\mu \in \mathfrak{g}^*$ in Patrick [1992, 1995].

The role of phases. One of the key things in the energy-momentum method is to keep track of which group drifts are possible. We discuss some basic examples of phases below. This is very important for the reconstruction process and for understanding the Hannay-Berry phase in the context of reduction (see Marsden, Montgomery and Ratiu [1990] and references therein). Noncompact groups come up in a number of examples, such as the dynamics of rigid bodies in fluids (underwater vehicles), which we discuss below.

Geometric phases or holonomy is also a useful concept in many *locomotion* problems as we will explain later. The mathematical foundations for phases in holonomic systems uses the theory of reduction, both Hamiltonian and Lagrangian; while for nonholonomic systems this foundation has been laid by Bloch, Krishnaprasad, Marsden, and Murray [1996], Ostrowski [1996] and many others. For the study of controllability and gaits, this setting of Lagrangian reduction has been very useful. An overview may be found in Lecture 3 and in Marsden and Ostrowski [1998].

In the underwater vehicle problem, the phase drifts one gets are interesting and important. They can be understood using the general machinery of *reduction by stages*. We return to this topic later in this lecture.

A related problem is the *orbit transfer problem* in which one tries to find a control to make a transition from one relative equilibrium to another using the natural dynamics as much as possible—for example, homoclinic connections provide *dynamical channels*. This is how some very interesting trajectories are generated for spacecraft mission planning.

Usefulness for pde. Arnold [1966c] was the *first* to prove a nonlinear stability theorem (the nonlinear Rayleigh inflection point criterion) for the two dimensional Euler equations of an ideal fluid, so these techniques are clearly effective. For an exposition of this and many related references, see Arnold and Khesin [1997]. Many other pde problems have been done as well, cf. Holm, Marsden, Ratiu and Weinstein [1985] for instance, as well as Lewis [1989].

The Benjamin-Bona theorem (Benjamin [1972], Bona [1975]) on stability of solitons for the KdV equation can be viewed as an instance of the energy-momentum method, including all the pde technicalities. See also Maddocks and Sachs [1993], and for example, Oh [1987], Grillakis, Shatah and Strauss [1987], although of course there are many subtleties special to the pde context.

Hamiltonian bifurcations. The energy-momentum method has also been used in the context of Hamiltonian bifurcation problems. One such context is that of free boundary problems, building on the work of Lewis, Montgomery, Marsden and Ratiu [1986] which gives a Hamiltonian structure for dynamic free boundary problems (surface waves, liquid drops, etc), generalizing Hamiltonian structures found by Zakharov. Along with the Arnold method itself, this is used for a study of the bifurcations of such problems in Lewis, Marsden and Ratiu [1987], Lewis, [1989, 1992a], Kruse, Marsden, and Scheurle [1993] and other references cited therein.

Converse to the energy-momentum method. Because of the block structure mentioned, it has also been possible to prove a “converse” of the energy-momentum method. The idea is to show that, if the second variation is indefinite, then the system is, in a sense we shall explain, unstable. One cannot, of course hope to do this literally since there are many systems (eg, examples studied by Chetayev) which are formally unstable, and yet their linearizations have eigenvalues lying on the imaginary axis. Most of these are presumably unstable, but this is a very delicate situation to prove analytically. Instead, the converse shows something easier but probably more important: with the addition of dissipation, the system is destabilized. This idea of *dissipation induced instability* goes back to Thomson and Tait in the last century. In the context of the energy-momentum method, Bloch, Krishnaprasad, Marsden and Ratiu [1994,1996] show that with the addition of appropriate dissipation, the indefiniteness of the second variation is sufficient to induce linear instability in the problem.

There are related eigenvalue movement formulas (going back to Krein) that are used to study non-Hamiltonian perturbations of Hamiltonian normal forms in Kirk, Marsden and Silber [1996]. There are interesting analogs of this for reversible systems in O’Reilly, Malhotra and Namamchivaya [1996]. These works contain citations to many more interesting references on these problems.

Extension of the energy-momentum method to nonholonomic systems. The energy-momentum method extends to certain nonholonomic systems. Building on the work on nonholonomic systems in Arnold, Kozlov, and Neishtadt [1988], Bates and Sniatycki [1993] and Bloch, Krishnaprasad, Marsden and Murray [1996], on the example of the Routh problem in Zenkov [1995] and on the vast Russian literature in this area, Zenkov, Bloch and Marsden [1997] show that there is a generalization to this setting. The method is effective in the sense that it applies to a wide variety of interesting examples, such as the rolling disk and a three wheeled vehicle known as the the roller racer. We will look at the geometry of nonholonomic systems in Lecture 3.

Relative equilibria and the underwater vehicle. Now we turn to stability (and bifurcation) of relative equilibria for mechanical systems with symmetry, using the dynamics of an underwater vehicle as a main example.

The main reference for the underwater vehicle example is Leonard and Marsden [1997] although the reader will probably want to consult Leonard [1997] and Holmes, Jenkins and Leonard [1997], as well as other references cited therein. A detailed understanding of the

underlying dynamics helps in control design; eg, when controls are limited, it helps to know if the underlying dynamics develops a nonlinear oscillation (Hopf bifurcation, flutter) or other instabilities as system parameters are varied.

Interesting Issues. One can handle certain nongeneric values of the momentum (where some degeneracy, or coincidence occurs) of the total (linear and angular) momentum using methods of Patrick [1992] and combining the energy-momentum theorems mentioned earlier with the technique of reduction by stages.

As indicated in the general introduction, in many examples, such as solitary waves, one has stability only modulo translations: nearby waves can move with different velocities and so drift. Similarly, with the underwater vehicle, one can have translational and rotational drift. The rotational drift is not arbitrary and can only happen around an axis that stays close to the original axis of spin of the relative equilibrium.

A key problem for the underwater vehicle is that the (symmetry) isotropy subgroup of the momentum is *noncompact*. Correspondingly, one can get interesting and perhaps unexpected phase drift—position and rotational drifting in the case of the underwater vehicle (that any control scheme must take into account). Leonard and Marsden [1997] generalize Patrick’s result, allowing some noncompactness, but taking the stability modulo a larger group. As we mention the main idea is to use reduction by stages, which we recall later, and apply the theorem of Patrick to the stage involving compactness. The method is implemented with the assistance of the energy-Casimir method. We will give specific examples of the outcome shortly.

Sample results for the underwater vehicle. For non-aggressive maneuvers, an accurate dynamical model for the underwater vehicle is by Kirchhoff’s equations: a rigid body in ideal potential flow. One instance of these equations were discussed in Lecture 1. Here we add the additional effect of *buoyancy*; *i.e.*, the center of mass and the center of buoyancy need not coincide. We will discuss the full dynamical equations shortly.

These equations are Lie-Poisson on the dual of a certain (semidirect product) Lie algebra. This structure is obtained in a natural way from T^*Q where $Q = SE(3)$ is the configuration space for the vehicle using *reduction theory for mechanical systems with symmetry*.

Notation and assumptions.

- neutrally buoyant and ellipsoidal vehicle
- centers of gravity and buoyancy need not be coincident
- inertia matrix of the body-fluid system: $J = \text{diag}(I_1, I_2, I_3)$
- mass matrix $M = \text{diag}(m_1, m_2, m_3)$ (J and M include the “added” inertias and masses due to the fluid)
- the mass of the body alone is m , and the acceleration of gravity is g
- $l =$ distance between center of buoyancy and center of gravity.

Relative equilibrium 1. In the first relative equilibrium we look at, the energy momentum technique reproduces a nonlinear extension of a classical result that may be found in Lamb [1932]. We assume the following:

- the vehicle is symmetric about the third (principal) axis

- it has coincident centers of buoyancy and gravity
- it translates with momentum (impulse) P_3^0 along the third axis and rotates with angular momentum (impulse) Π_3^0 about the same axis (see Figure 2.2).

The techniques outlined above applied to this case show that: *This motion is stable modulo rotations about the third axis and translations in any direction provided that*

$$\left(\frac{\Pi_3^0}{P_3^0}\right)^2 > 4I_2 \left(\frac{1}{m_3} - \frac{1}{m_2}\right).$$

Roughly speaking, this means that “blunt” motion is stable, while “streamline” and slowly spinning motion is unstable.

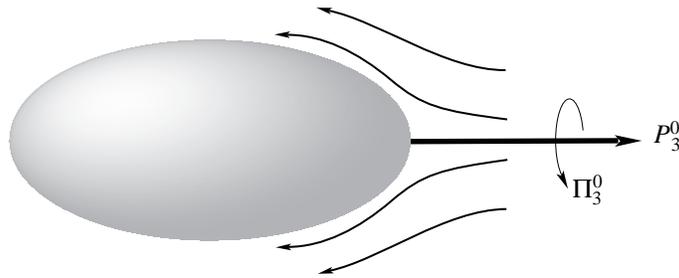


Figure 2.2: Translating and spinning ellipsoid.

Relative equilibrium 2. The second sample relative equilibrium has the following features:

- The vehicle has noncoincident centers of buoyancy and gravity oriented with the third principal axis parallel to the direction of gravity
- It is translating (not spinning) with momentum P_2^0 along the second principal axis
- It need not be symmetric (see Figure 2.3).

This motion is stable modulo translations in any direction provided that the system is bottom heavy ($l > 0$) and

$$m_2 > m_1, \quad mgl > \left(\frac{1}{m_2} - \frac{1}{m_3}\right) (P_2^0)^2.$$

Relative equilibrium 3. The third example has a nongeneric value of the momentum and is the same as Case 1 except that there is no translation; *i.e.*, $P_3^0 = 0$.

This motion is stable modulo rotations about the third axis and translations in any direction provided that I_3 is largest or smallest of the moments of inertia. (Unlike the free rigid body, this need not mean stability of rotations about the major or minor axis of the body.)

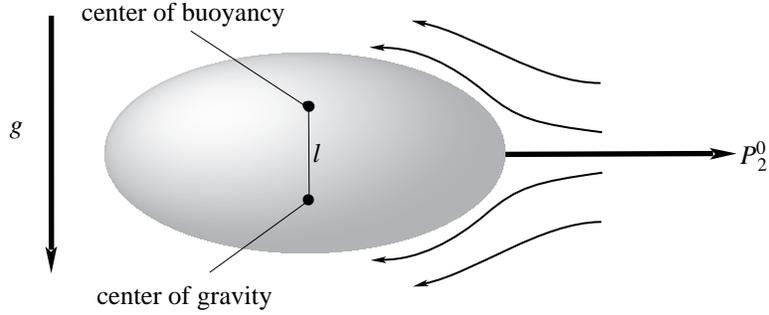


Figure 2.3: Adding buoyancy.

Relative equilibrium 4. This case is also nongeneric but for a different reason; now the direction of translation is parallel to direction of gravity. Here,

- the centers of buoyancy and gravity are noncoincident and are oriented with the third principal axis parallel to the direction of gravity
- the vehicle is symmetric about the third axis
- the vehicle is rising (or falling) with momentum P_3^0 along the third axis rotates with angular momentum Π_3^0 . See Figure 2.4.

This motion is stable modulo rotations about the third axis and translations in any direction provided that

$$mgl > \left(\frac{1}{m_3} - \frac{1}{m_2} \right) (P_3^0)^2 - \frac{\bar{a}_2}{4} (\Pi_3^0)^2,$$

where

$$\bar{a}_2 = \frac{m_1}{m_1 I_2 - m^2 l^2}.$$

There is an interesting bifurcation in relative equilibrium 4: a Hamiltonian Krein-Hopf bifurcation (the splitting of eigenvalues off the imaginary axis) and the passing (resonance) of eigenvalues. We will return to this shortly.

Stability theorem for nongeneric momenta. Now we give a few of the details on the method and how it works. The setting is a phase space P (a symplectic manifold) with symplectic form Ω . Assume as before, that we have a Lie group G with a Lie algebra \mathfrak{g} and that G acts symplectically on P with an equivariant momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$. We also have a G -invariant Hamiltonian H that describes the dynamics of interest. The associated Hamiltonian vector field is denoted X_H and the dynamical equations of motion are $\dot{z} = X_H(z)$. Noether's theorem of course states that \mathbf{J} is a constant of the motion.

Relative equilibria and energy-momentum method. Consider a *relative equilibrium* $z_e \in P$: there is a Lie algebra element $\xi_e \in \mathfrak{g}$ such that $\exp(\xi_e t)z_e$ is a dynamic orbit. Let $\mu_e = \mathbf{J}(z_e)$. Assume that z_e is a regular point of \mathbf{J} (equivalent to no infinitesimal symmetries). We need *not* assume that μ_e is a generic point. As mentioned earlier the *augmented energy* $H_{\xi_e} = H - \langle J, \xi_e \rangle$ has a critical point at z_e . Next, calculate the second

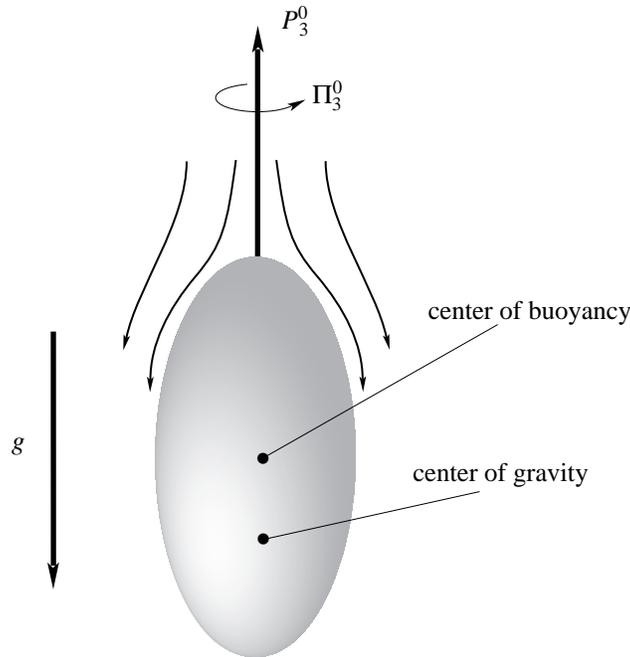


Figure 2.4: A rising, spinning vehicle.

derivative $\delta^2 H_{\xi_e}(z_e)$ of H_{ξ} at z_e . Note that $\ker \mathbf{DJ}(z_e)$ is the tangent space to the level set $\mathbf{J}(z) = \mu_e$ at the point z_e . The tangent space to the G_{μ_e} -orbit of z_e at z_e is given by the vector space consisting of infinitesimal generators $\xi_P(z_e)$ as ξ ranges over \mathfrak{g}_{μ_e} , the Lie algebra of G_{μ_e} .

Now choose a vector subspace $E_{z_e} \subset \ker \mathbf{DJ}(z_e)$ that complements the tangent space to the G_{μ_e} -orbit of z_e ; *i.e.*, $\ker \mathbf{DJ}(z_e) = E_{z_e} \oplus T_{z_e} G_{\mu_e}(z_e)$, as in Figure 2.1.

Assume that the second derivative of $\delta^2 H_{\xi_e}(z_e)$ restricted to E_{z_e} is definite.

Patrick's theorem. Assume the restriction of the coadjoint action of G_{μ_e} on \mathfrak{g}^* is proper and there is an inner product on \mathfrak{g}^* that is invariant under this action; for example, these both hold if G_{μ_e} is compact (and this in turn holds if G is compact). *Then the relative equilibrium z_e is stable modulo G_{μ_e} .*

There is a geometric phase drift along the group directions; *i.e.*, in the G_{μ_e} directions that are not controlled by the stability theorem.

Relation to energy-Casimir method. If $P = T^*G$, then the dynamics reduces to Lie-Poisson dynamics on \mathfrak{g}^* relative to the reduced Hamiltonian h . One can check the preceding hypothesis on the second variation by finding a function C on \mathfrak{g}^* and constant on the coadjoint orbit through the relative equilibrium for which the second variation of $d^2(h + C)$ is definite. In the nongeneric case we use *subcasimir as well as Casimir functions*.

Extra drift. In the underwater vehicle example, the properness assumption fails. In fact, simulations for the underwater vehicle show that one really has a new phenomenon of drift and so stability has to be taken modulo a larger group than G_{μ_e} .

To illustrate computationally what happens, consider special initial conditions near the equilibrium in Case 1:

- specific inertia and mass matrix parameters are $I_1 = I_2 = 1$, $I_3 = 0.5$, $m_1 = m_2 = 1$, $m_3 = 0.8$
- equilibrium momentum: $\Pi_3^0 = I_3\Omega_3^0$ and $P_3^0 = m_3v_3^0$
- equilibrium values of the velocities are $\Omega_3^0 = 2.5$ and $v_3^0 = 1$
- the stability results stated above *show that this is a stable equilibrium modulo drift in γ_3 (angle about the third axis) and b (the translational position)*
- the initial conditions chosen are: $\Omega(0) = (0.01, -0.008, 2.5)^T$, $v(0) = (0.009, 0.011, 1)^T$, $\gamma(0) = 0$ and $b(0) = 0$.

Figures 2.5 and 2.6 show the stability in the velocities and the drift in the angle γ_3 and the translational position b .

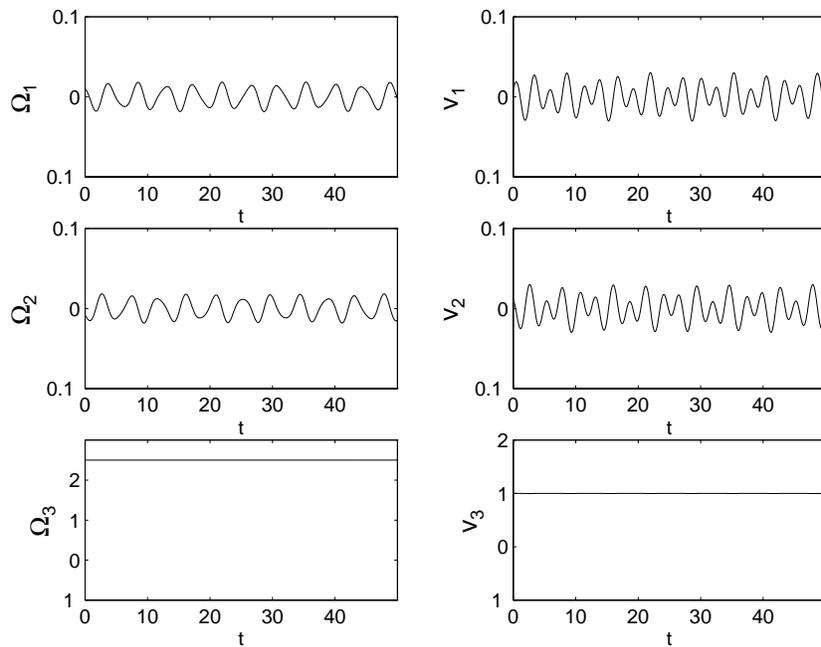
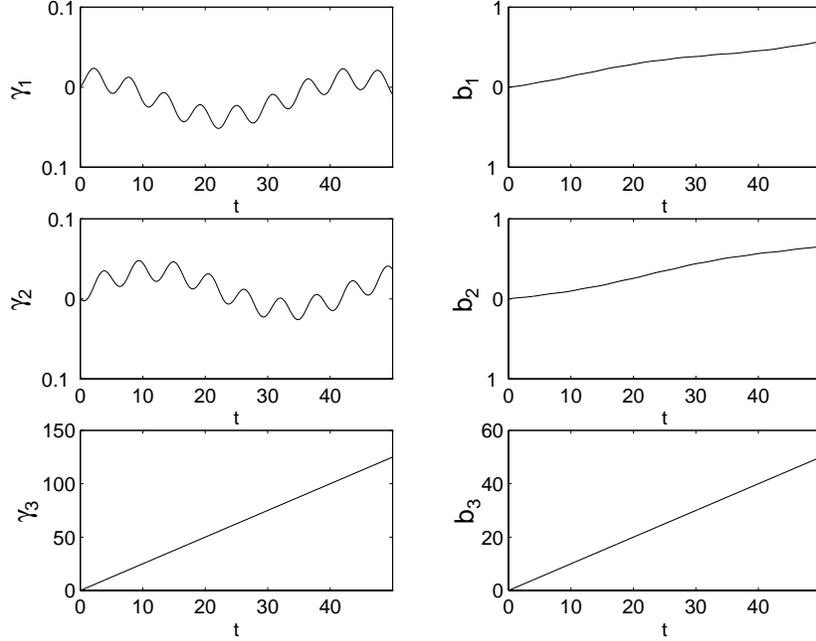


Figure 2.5: The velocities are stable.

Reduction by stages. In the first lecture we discussed some generalities on reduction by stages. Now we discuss the specific case of semidirect products, which is the case of interest for the underwater vehicle.

Start with a Lie group that is a semidirect product, $S = G \ltimes V$ where V is a vector space and the Lie group G acts on V (and hence on its dual space V^*). The Lie algebra of S is denoted $\mathfrak{s} = \mathfrak{g} \ltimes V$.

Assume we have a symplectic action of S on a symplectic manifold P and that there is an equivariant momentum map $\mathbf{J}_S : P \rightarrow \mathfrak{s}^*$. The translation subgroup V thus also acts on

Figure 2.6: The angle γ_3 and the position \mathbf{b} drift.

P and has a momentum map $\mathbf{J}_V : P \rightarrow V^*$ obtained from \mathbf{J}_S by projection from \mathfrak{s}^* onto V^* .

First reduce P by V at the value $a \in V^*$ (assume it to be a regular value) to get the **first reduced space** $P_a = \mathbf{J}_V^{-1}(a)/V$. Next, form the group G_a consisting of elements of G that leave the point $a \in V^*$ fixed using the action of G on V^* . Then one proves (by not entirely obvious definition chasing) that G_a acts symplectically on P_a and has an induced equivariant momentum map $\mathbf{J}_a : P_a \rightarrow \mathfrak{g}_a^*$, where \mathfrak{g}_a is the Lie algebra of G_a . Then reduce P_a at the point $\mu_a := \mu|_{\mathfrak{g}_a} \in \mathfrak{g}_a^*$ to get the **second reduced space** $(P_a)_{\mu_a} = \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

Theorem 2.1 (Reduction by stages for semidirect products.) *The symplectic reduced space $(P_a)_{\mu_a}$ is symplectically diffeomorphic to the reduced space P_σ obtained by reducing P by S at the point $\sigma = (\mu, a)$.*

From this we get the well known result of Ratiu, Guillemin, Sternberg, Marsden and Weinstein (see Marsden, Ratiu and Weinstein [1984a,b] and references therein) as a special case:

Corollary 2.2 (Semidirect product reduction.) *The reduction of T^*G by G_a at values $\mu_a = \mu|_{\mathfrak{g}_a}$ gives a space that is isomorphic to the coadjoint orbit through the point $\sigma = (\mu, a) \in \mathfrak{s}^* = \mathfrak{g}^* \times V^*$, the dual of the Lie algebra \mathfrak{s} of S .*

Of course if one has an S -invariant Hamiltonian H , one can reduce the dynamics of a given H in two stages.

This theorem explains why one gets Lie-Poisson dynamics for the underwater vehicle system on the dual of the Lie algebra of the semidirect product $SO(3) \ltimes (\mathbb{R}^3 \times \mathbb{R}^3)$, starting with the physical configuration space $Q = SE(3)$ and reducing it by the symmetry group $SE(2) \times \mathbb{R}^2$. This remarkable fact can also be checked directly from the dynamical equations which we write down shortly.

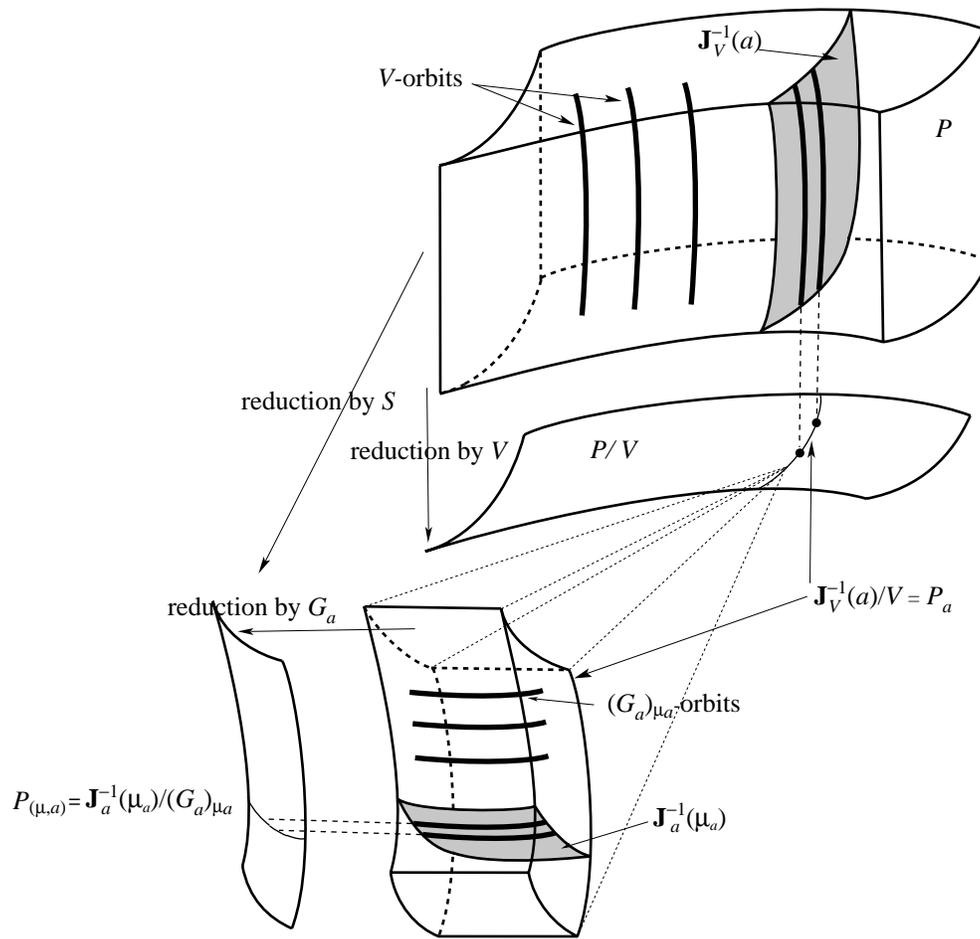


Figure 2.7: Reduction by stages.

Generalized stability theorem. To deal with the non-compactness, we make some special assumptions. The action of $(G_a)_{\mu_a}$ on \mathfrak{g}_a^* is proper and there is an inner product on \mathfrak{g}_a^* that is invariant under this action; for example, both of these conditions hold if $(G_a)_{\mu_a}$ is compact (and this holds if G is compact).

Choose a vector subspace $E_{[z_e]} \subset \ker \mathbf{D}J_a([z_e])$ that complements the tangent space to the $(G_a)_{\mu_a}$ -orbit of $[z_e]$. Let the Hamiltonian reduced to the first stage space P_a be denoted H_a and assume that the second derivative of $H_a - (J_a)_{\xi_e}$ at $[z_e] \in P_a$ restricted to $E_{[z_e]}$ is definite.

Theorem 2.3 *In addition to the above, assume $\mu_a \in \mathfrak{g}_a^*$ is generic. Then the point $[z_e] \in P_a$ is Liapunov stable modulo the action of $(G_a)_{\mu_a}$ for the dynamics of H_a and z_e is Liapunov stable in P modulo the action of $(G_a)_{\mu_a} \otimes V$.*

One can verify the hypothesis on the second variation by using an argument on the Poisson reduced space P/S as in the energy-Casimir and energy-subcasimir method.

In Leonard and Marsden [1997] a more delicate energy criterion is given for non-generic μ_a which is relevant for and applied to the rising, spinning vehicle.

Dynamics of an underwater vehicle. Now we give a few more details on why the dynamics of the underwater vehicle problem can be viewed as Lie-Poisson dynamics; *i.e.*, as generalized Euler equations.

Assume the underwater vehicle is a neutrally buoyant, rigid body submerged in an infinitely large volume of incompressible, inviscid, irrotational fluid which is at rest at infinity. The dynamics of the body-fluid system are described by Kirchhoff's equations, where we assume the only external forces and torques acting on the system are due to buoyancy and gravity.

Consider the group W , the semidirect product of $SO(3)$ with *two* copies of \mathbb{R}^3 , *i.e.*, $W = SO(3) \ltimes (\mathbb{R}^3 \times \mathbb{R}^3)$, where we take the action of $SO(3)$ on $\mathbb{R}^3 \times \mathbb{R}^3$ to be $R \cdot (b, w) = (Rb, Rw)$. The group multiplication in W is

$$(R, b, w)(R', b', w') = (RR', Rb' + b, Rw' + w),$$

while the Lie algebra of W is $\mathfrak{w} = \mathfrak{so}(3) \ltimes \mathbb{R}^3 \times \mathbb{R}^3$.

Let $(\Pi, P, \Gamma) \in \mathfrak{w}^*$ correspond, to angular and linear components of the *impulse* of the system and the direction of gravity in body coordinates. The Poisson bracket used is the Lie-Poisson bracket on \mathfrak{w}^* . Let Ω and v , respectively, be the angular and translational velocity of the vehicle in body coordinates. From the Lagrangian given in Lecture 1, we get

$$\begin{aligned} \Pi &= J\Omega + Dv \\ P &= Mv + D^T\Omega. \end{aligned}$$

Here, J is the sum of the body inertia matrix plus the *added inertia matrix* and M is the mass matrix for the body alone, plus the *added mass matrix* associated with the fluid.

For an ellipsoidal body, with appropriate choice of body-fixed coordinate frame, M and J are diagonal and $D = m\hat{r}_G$, where r_G is the vector from the center of buoyancy to the center of gravity. This simplifies matters somewhat.

The underwater vehicle dynamics has Lie-Poisson form on \mathfrak{w}^* . This fact follows by constructing the physical phase space of $T^*SE(3)$ reduced by the symmetry group $SE(2) \times \mathbb{R}$ with the aid of the semidirect product reduction theory. The Hamiltonian comes from the Lagrangian in Lecture 1 via the Legendre transform with the addition of a buoyancy potential:

$$h(\Pi, P, \Gamma) = \frac{1}{2}(\Pi^T A \Pi + 2\Pi^T B^T P + P^T C P - 2mg(\Gamma \cdot r_G)), \quad (2.1)$$

where we determine A, B, C by inverting the block matrix giving the metric (or mass matrix):

$$\begin{aligned} A &= (J - DM^{-1}D^T)^{-1} \\ B &= -CD^T J^{-1} = -M^{-1}D^T A \\ C &= (M - D^T J^{-1}D)^{-1}. \end{aligned}$$

The Lie-Poisson equations of motion are

$$\begin{aligned} \dot{\Pi} &= \Pi \times \Omega + P \times v - mg\Gamma \times r_G \\ \dot{P} &= P \times \Omega \\ \dot{\Gamma} &= \Gamma \times \Omega. \end{aligned}$$

Recall that a point $\sigma = (\Pi, P, \Gamma) \in \mathfrak{w}^*$ is called **generic** if the coadjoint orbit through that point has maximal dimension six. These are three independent **Casimir functions** (functions invariant under the coadjoint action):

$$C_1(\Pi, P, \Gamma) = P \cdot \Gamma, \quad C_2(\Pi, P, \Gamma) = \|P\|^2, \quad C_3(\Pi, P, \Gamma) = \|\Gamma\|^2.$$

A point $\sigma = (\Pi, P, \Gamma) \in \mathfrak{w}^*$ is nongeneric if the vectors P and Γ are not parallel. If P and Γ are not both zero, the coadjoint orbit through σ has dimension four.

Besides the three Casimirs defined above, two additional conserved quantities (called *subcasimirs*) on the nongeneric coadjoint orbits are

$$C_4(\Pi, P, \Gamma) = \Pi \cdot P, \quad C_5(\Pi, P, \Gamma) = \Pi \cdot \Gamma.$$

If $P = \Gamma = 0$ with $\Pi \neq 0$, then the coadjoint orbit through ω has dimension two. An additional subcasimir on this nongeneric coadjoint orbit is

$$C_6(\Pi, P, \Gamma) = \|\Pi\|^2.$$

Continuing along these lines with the energy-momentum and energy-Casimir method gives the stability conditions as stated earlier; we refer to Leonard [1997] and Leonard and Marsden [1997] for details.

Hamiltonian Hopf bifurcation. Consider now the fourth relative equilibrium. The eigenvalues of the linearization of the dynamics at the equilibrium as the equilibrium linear momentum P_3^0 is varied, are shown in Figure 2.8.

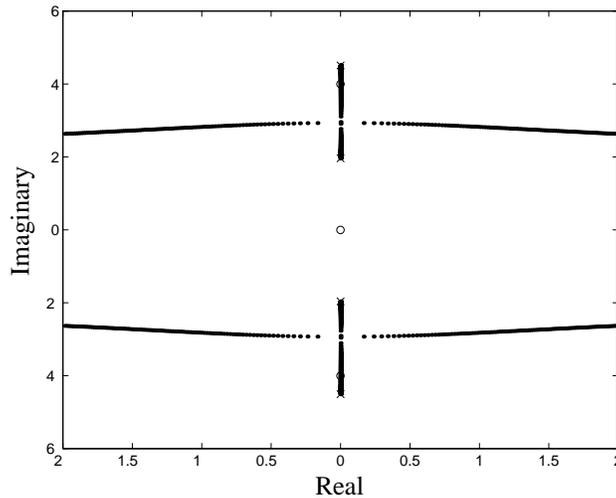


Figure 2.8: Eigenvalue movement for relative equilibrium 4 as the vertical momentum is increased.

There are three eigenvalues fixed at the origin and two eigenvalues fixed at $\pm(\Pi_3^0/I_3)i$. The remaining four eigenvalues move as the parameter P_3^0 is varied. These eigenvalues are on the imaginary axis as long as the stability condition is met. Assuming $m_3 < m_2$ (*i.e.*, $l_3 > l_2$) and $mgl > 0$, as P_3^0 is increased, the pair of eigenvalues above the real axis and the pair below each meet and then split off the imaginary axis—saddle points develop. The point at which each pair of eigenvalues meets, *i.e.*, the Hamiltonian-Krein-Hopf bifurcation point, corresponds to the value of P_3^0 that makes the stability condition an equality. Thus, the stability analysis accurately predicts the bifurcation point. Before the bifurcation occurs, *i.e.*, while the eigenvalues are all on the imaginary axis, each of the eigenvalues fixed at $\pm(\Pi_3^0/I_3)i$ is passed by one of the moving eigenvalues. This passing

(resonance) of eigenvalues occurs when P_3^0 is such that

$$mgl = \left(\frac{1}{m_3} - \frac{1}{m_2} \right) (P_3^0)^2.$$

Remarkably, this passing event may be precisely detected by the more delicate stability criterion (at non-generic points) that was mentioned earlier. In this computation, the inertia and mass matrix parameters were chosen (for illustration only) to be $I_1 = I_2 = 1$, $I_3 = 0.5$, $m_1 = m_2 = 1$, $m_3 = 0.8$ and $m = 0.5$, $g = 9.8$, $l = 0.5$. The equilibrium angular momentum is $\Pi_3^0 = 2$ and P_3^0 is increased from 1 to 6.

The five eigenvalues that remain fixed are drawn as circles (note there are three eigenvalues at the origin). The crosses indicate the positions of the four remaining eigenvalues at the point when $P_3^0 = 1$. The dotted lines show the paths of these four eigenvalues as P_3^0 is increased to 6. The eigenvalue crossing occurs when $P_3^0 = \sqrt{9.8} \approx 3.13$ and the Hamiltonian Hopf bifurcation point corresponds to $P_3^0 \approx 3.75$.

It would be interesting to study the Hamiltonian-Krein-Hopf bifurcation in more detail. (Be warned that standard versions of this theorem will not work.) It would also study the effects of symmetry breaking (such as the S^1 symmetry of the vehicle in the case of the rising vehicle) needs additional attention. The techniques of Knobloch, Mahalov and Marsden [1994] may prove useful in this regard. A complication is the nongeneric nature of the coadjoint orbit, so the theory of eigenvalue movement (see Dellnitz, Melbourne and Marsden [1992]) requires additional work.

Some feedback stabilization methods (Bloch, Krishnaprasad, Marsden and Sanchez, [1992]) and Bloch, Leonard and Marsden [1997] may be applied to stabilize otherwise unstable equilibria. This is an interesting direction for future research; the basic ideas of the method are given in Lecture 4.

More realistic models of the fluid dynamics, especially in cases when vorticity is generated by the body-fluid interaction as well as the inclusion of elastic and flexible properties of the body would also be interesting, so would the effect of dissipation after Bloch, Krishnaprasad, Marsden and Ratiu [1994, 1996].

The underwater vehicle problem should continue to provide interesting additional motivations for the continued development of the basic theory.

The planar skater and phases. We now turn our attention to the topic of geometric phases, which we have already seen is important in the underwater vehicle. We begin with a closer look at the basic theory.

For a mechanical system with Hamiltonian of the form kinetic plus potential and with a symmetry, one can think of the zero momentum map level set as a **constraint**. Let g denote the **group position** and r the **internal shape**: then we can write the equation $J = 0$ as

$$g^{-1}\dot{g} = -A(r)\dot{r},$$

which defines A , the **mechanical connection**; this definition agrees with the one we had earlier. For nonzero momentum values μ of J , then the equation $J = \mu$ can be written as

$$g^{-1}\dot{g} = -A(r)\dot{r} + (I(r, g))^{-1}\mu,$$

which also defines I , the **locked inertia tensor**.

A good example illustrating these ideas is the planar skater, shown in Figure 2.9. It consists of three linked rigid bodies in the plane (imagine them sliding on an ice rink).

Suppose that the angles ψ_1 and ψ_2 perform a cyclic or periodic motion (imagine motors in the joints). Note that if the angular momentum starts out zero, it remains so since no

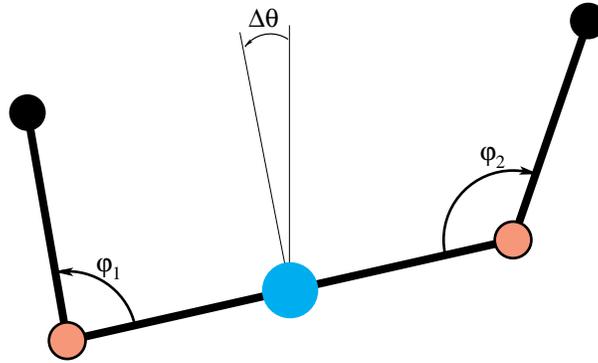


Figure 2.9: The planar skater consists of three interconnected bodies that are free to rotate about their joints.

external torques act on the system. The cyclic inputs to this system are shown as a base input curve, while the actual trajectory of the motion is shown lifted above the input curve.

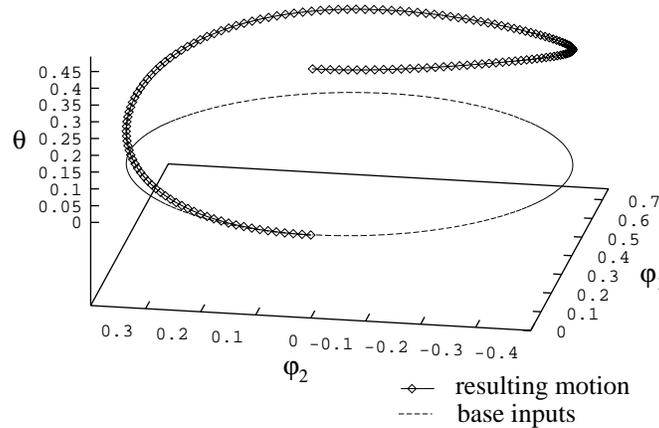


Figure 2.10: Input and output motions for the planar skater.

After completing one cycle of internal shape changes, the skater has undergone a net rotation (change in θ). The area enclosed by the base inputs is proportional to the net rotation. This is a simple example of the *geometric phase*, or *holonomy*, associated with the cyclic shape inputs. The planar skater is a simple enough example that it can be dealt with by hand in a simple way, using the definition of angular momentum and its conservation. More sophisticated examples, such as the falling cat or the reorientation of spacecraft requires more elaborate mathematics. One of the keys to a deeper understanding is to get a better vision of holonomy.

Holonomy is also a basic notion in geometry per se and can be illustrated using parallel transport on the sphere (the same geometry is used to understand the Foucault pendulum). See Figure 2.11. In this figure one parallel transports an arrow around a triangle on the sphere whose sides are made up of portions of great circles. In this case, parallel transport along the arcs is parallel transport in the naive sense. Notice that when the arrow returns to its original location, it has rotated relative to its original position. This rotation is another

example of a *phase shift*.

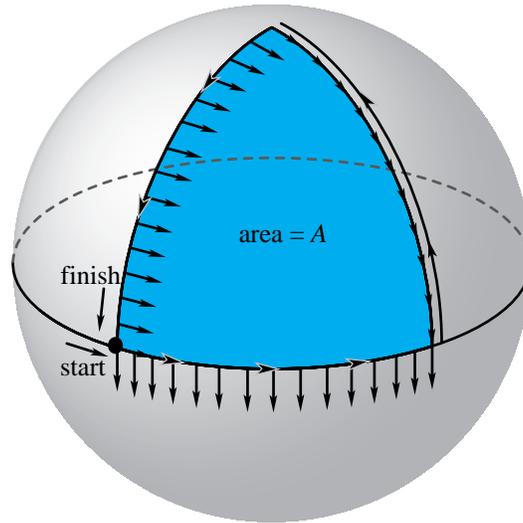


Figure 2.11: A parallel movement of a vector around a spherical triangle produces a phase shift.

The *general* notion of *holonomy* in geometry involves the splitting of the tangent space of a bundle with the use of a connection (more generally a distribution), as shown in Figure 2.12.

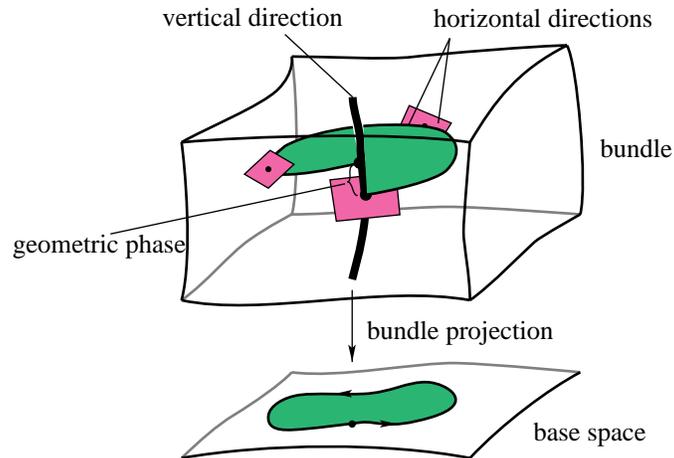


Figure 2.12: A connection divides the space into vertical and horizontal directions.

Interestingly, geometric phases or Berry phases in *quantum mechanics* can be viewed as a special case of phases for classical systems by using the (now well known) fact that with Hilbert space as our phase space (the symplectic structure being the imaginary part of the Hermitian inner product), we get the Schrödinger equation as a *special case of Hamilton's equations*. The symmetry group is S^1 phase shift and the reduction is the classical process of relating quantum mechanics on Hilbert space to that on projective Hilbert space studied

by Bargmann and Wigner. (See also Marsden, Montgomery and Ratiu [1990] and Marsden and Ratiu [1998] for a discussion.)

Phases for a relative equilibrium for the underwater vehicle were shown in earlier figures. The **body angular and linear velocities remain stable** while the group variables (rotation around one axis and translations) can drift, as we have explained.

Phases for the three-wave interaction. We return now to the three wave interaction discussed in the first lecture and briefly discuss phases for the problem. We now want to **reconstruct** the dynamics of the original system on \mathbb{C}^3 from that on the three wave surface.

A standard direct approach would be to proceed as follows. First, use (two of) the constants of motion K_1, K_2, K_3 along with (one of) the reduced coordinates Z_1 and Z_2 to determine the amplitudes $|q_1|^2, |q_2|^2, |q_3|^3$. Then, write $q_k = |q_k| \exp(i\phi_k)$ and substitute into the equations of motion; this yields evolution equations for the phases ϕ_k which may then be integrated. While the amplitudes undergo periodic orbits for most initial data on the reduced phase space (other than the homoclinic orbits), the phases associated with them obtain a shift, which is then implicitly obtained via quadratures.

A method that is more geometric in nature, is based on Montgomery's [1991b] derivation of the phase formula for the rigid body and that for the polarization phase shift. The general theory of this is found in Marsden, Montgomery and Ratiu [1990]. Let $M \subset \mathbb{C}^3$ be the manifold defined by setting the conserved quantities to specific values (a level set of the momentum map). We will construct closed curves in M in two pieces as follows.

The first portion c_0 is simply the dynamical trajectory in M joining two points P_0 and P_1 . It covers a closed curve in the reduced space. The invariants define the **reduction map** $W : \mathbb{C}^3 \rightarrow \mathbb{R}^3$ (the image is the three-wave surface) so that the curve c_0 projects onto a closed trajectory in the base space under W . (See Figure 2.13.)

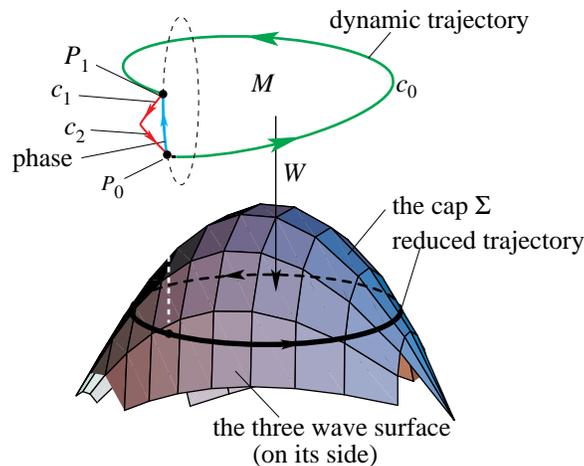


Figure 2.13: Curves used for the phase calculation.

In M we **introduce three curves** c_1, c_2 , and c_3 that close the curve by connecting the points P_1 and P_0 using the group actions: The curve c_1 goes with the first S^1 action, etc. Then $c_1 \cup c_2$ closes the curve, as shown in the figure, as do $c_1 \cup c_3$ and $c_2 \cup c_3$. Specifically, if $P_0 = (q_1, q_2, q_3)$ and if $P_1 = (q_1 \exp i\phi_1, q_2 \exp i(\phi_1 + \phi_2), q_3 \exp i\phi_2)$, we let c_1 be the opposite (run in the opposite sense) of the curve

$$c_1^{\text{opp}}(t) = (q_1 \exp it\phi_1, q_2 \exp it\phi_1, q_3),$$

$0 \leq t \leq 1$, which is the action by $t\phi_1$ for the first group action and let c_2 be the opposite of

$$c_2^{\text{OPP}}(t) = (q_1 \exp i\phi_1, q_2 \exp i(\phi_1 + t\phi_2), q_3 \exp it\phi_2)$$

which is the action by $t\phi_2$ for the second group action (These actions and the corresponding conserved quantities were defined in the first lecture).

Let Θ be a **canonical one form** on \mathbb{C}^3 , a scaling of the Poincaré one form defined by

$$\Theta = \frac{1}{2}(p_k dq^k - q^k dp_k).$$

Thus, $d\Theta = -\Omega$ is the symplectic form. By Stokes theorem,

$$\int_{c_0} \Theta + \int_{c_1} \Theta + \int_{c_2} \Theta = \int_{S_3} d\Theta,$$

where S_i are surfaces that project to the cap Σ shown in the figure. (A more careful argument—as in holonomy theorems—shows that the existence of these surfaces is not necessary.)

Noting that

$$\int_{S_i} \Omega = \int_{\Sigma} \Omega_r,$$

where Ω_r is the **reduced symplectic form** (the symplectic form—an area element—on the three wave surface), we find

$$\int_{c_0} \Theta + \int_{c_1} \Theta + \int_{c_2} \Theta = - \int_{\Sigma} \Omega_r.$$

On the **group curves** c_i we have

$$\int_{c_i} \Theta = -K_i \phi_i,$$

$i = 1, 2, 3$ where K_i is the (constant) conserved quantity and ϕ_i is the phase shift associated with the i th S^1 . This is because K_i is *homogeneous of degree 2*.

On the **dynamic trajectory** c_0 we have

$$\int_{c_0} \Theta = \int_{c_0} \langle \Theta, X_H \rangle dt = \frac{3}{2}HT$$

where H is the *constant energy* of the trajectory and T is the *period* of the reduced trajectory. This is because H is *homogeneous of degree 3* in the q_i .

Letting $A(\Sigma)$ denote the *symplectic area* of Σ and putting the above equations together, we obtain a *phase identity*:

$$\frac{3}{2}HT - \phi_1 K_1 - \phi_2 K_2 = -A(\Sigma).$$

An important special case occurs on the fixed point space of the interchange symmetry $q_1 \leftrightarrow q_3$ in the case that $\gamma_1 = \gamma_3$. This corresponds to the case of **second harmonic generation** in nonlinear optics. In this case, from the formulas for the two points P_1 and P_2 , we see that we can assume $\phi_1 = \phi_2 =: \phi$ and so we get an explicit formula for the phase:

$$\phi = \frac{3HT + 2A(\Sigma)}{2(K_1 + K_2)}.$$

More general formulas for the phase are computed in Alber, Luther, Marsden and Robbins [1998b].

These methods of analyzing the three wave interaction also allows one to apply control ideas to manipulate the phases (as in what are called cascaded nonlinearities); see Alber, Luther, Marsden and Robbins [1998a] for more information.

Lecture 3

Systems with Rolling Constraints and Locomotion

In this lecture we consider the basic mechanics of systems with rolling constraints and related questions of locomotion. Some of the exposition follows the development in Bloch, Krishnaprasad, Marsden and Murray [1996] and Marsden and Ostrowski [1998].

For systems with *rolling constraints* or more generally *nonholonomic systems*, one finds the equations of motion and properties of the solutions (such as the fate of conservation laws) using the Lagrange-d'Alembert principle. These systems are *not literally variational* but the basic mechanics still comes down to $F = ma$.

Consider a *configuration space* Q and a *distribution* \mathcal{D} that describes the kinematic constraints; \mathcal{D} is a collection of linear subspaces: $\mathcal{D}_q \subset T_qQ$, for $q \in Q$. We say that $q(t) \in Q$ *satisfies the constraints* if $\dot{q}(t) \in \mathcal{D}_{q(t)}$. This distribution is, in general, *nonintegrable* in the sense of Frobenius' theorem; *i.e.*, the constraints can be *nonholonomic*. Anholonomy is measured by the *curvature* of \mathcal{D} .

A simple example of a nonholonomic system is the *rolling disk* (see Figure 3.1). Here,

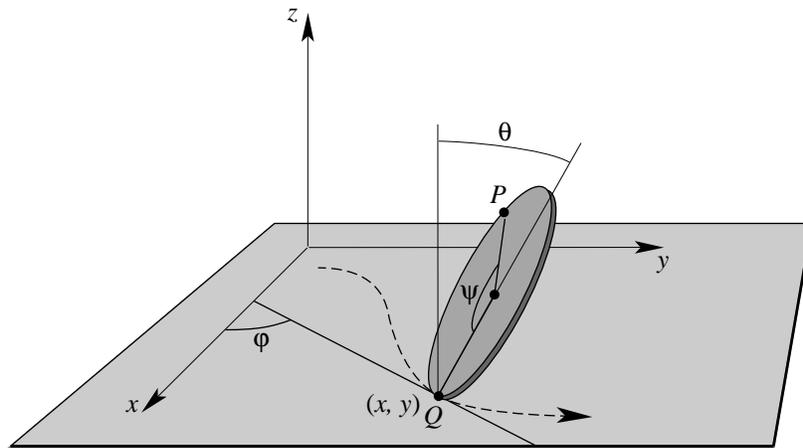


Figure 3.1: The rolling disk.

the constraints of rolling without slipping define the distribution \mathcal{D} :

$$\begin{aligned}\dot{x} &= -\dot{\psi}R \cos \phi \\ \dot{y} &= -\dot{\psi}R \sin \phi.\end{aligned}$$

These equations for the velocities may be interpreted as defining a linear subspace of the tangent space to the configuration space. These linear spaces then make up the constraint distribution \mathcal{D} .

The Lagrange-d'Alembert principle. The system dynamics is determined by a *Lagrangian* $L : TQ \rightarrow \mathbb{R}$, usually the kinetic minus the potential energy. The basic equations of motion are given by requiring that $q(t)$ satisfy the constraints *and* that

$$\delta \int_a^b L(q, \dot{q}) dt = 0,$$

for all variations satisfying $\delta q(t) \in \mathcal{D}_q(t)$, $a \leq t \leq b$.

Consistent with the fact that each \mathcal{D}_q is a *linear* subspace, we consider only homogeneous velocity constraints. The extension to affine constraints is straightforward, as in Bloch, Krishnaprasad, Marsden and Murray [1996].

This is *not a variational principle in the usual sense* (this issue of variational or not was discussed extensively and “put to rest” by Korteweg [1899]). Recall that one must also put constraints on the variations for the *Euler-Poincaré equations* as we saw in Lecture 1.

Standard arguments in the calculus of variations show that this “constrained variational principle” is equivalent to the equations

$$-\delta L := \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i = 0, \quad (3.1)$$

for all variations δq such that $\delta q \in \mathcal{D}_q$ at each point of the underlying curve $q(t)$. These equations are equivalently written as

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_i,$$

where λ_i is a set of Lagrange multipliers ($i = 1, \dots, n$), representing the force of constraint. Intrinsically, this multiplier λ is a section of the cotangent bundle over $q(t)$ that annihilates the constraint distribution. The Lagrange multipliers are often determined by using the condition that $\dot{q}(t)$ lies in the distribution.

The roller racer. Another example is the *roller racer*—it is a wheeled vehicle with two segments connected by a rotational joint, as in Figure 3.2.

The configuration space is given by $SE(2) \times S^1$, whose elements give the overall position and orientation of the vehicle plus the internal shape angle ϕ . The constraints are defined by the condition of rolling without slipping, as in the case of the falling penny.

The roller racer is interesting because it *generates locomotion*. *If you climb aboard and wiggle the joint, the vehicle moves!*¹ We will come back to this locomotion question shortly.

¹(See <http://www.isr.umd.edu/krishna/> for interesting movies of this.)

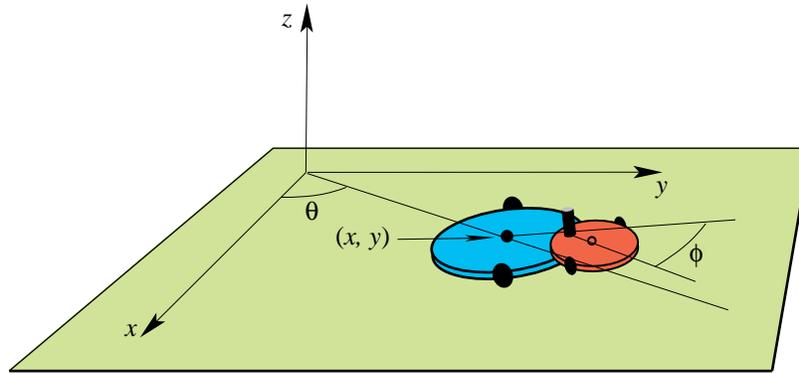


Figure 3.2: The roller racer—Tennessee racer.

The rattleback. This system, also called the *wobblestone*, is another much studied example, illustrating the *lack of conservation of angular momentum*. This is demonstrated by the reversal of its direction of rotation when spun. General theory provides a replacement, for the conservation law from “standard” mechanics, namely the *momentum equation*.

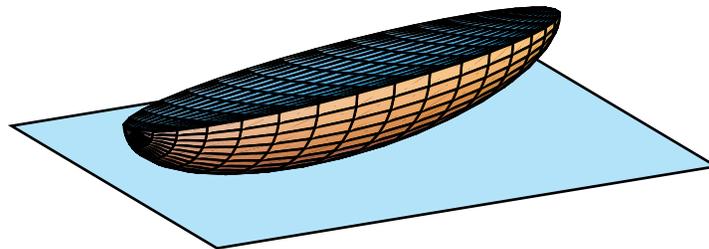


Figure 3.3: The rattleback.

Special features of nonholonomic mechanics. Here are some of the key features of nonholonomic systems that set them apart from “ordinary” mechanical systems:

- symmetry *need not lead to conservation laws*, but rather lead to an interesting *momentum equation*,
- equilibria can be stable, with *some* variables being *asymptotically stable*,
- energy is *still conserved*,
- Jacobi’s identity for Poisson brackets *can fail*.

To explore the structure of the Lagrange-d’Alembert equations in more detail, let $\{\omega^a\}$, $a = 1, \dots, k$ be a set of k independent one forms whose vanishing describes the constraints; i.e., the distribution \mathcal{D} . One can introduce local coordinates $q^i = (r^\alpha, s^a)$ where $\alpha = 1, \dots, n - k$, in which ω^a has the form

$$\omega^a(q) = ds^a + A_\alpha^a(r, s)dr^\alpha,$$

where the summation convention is in force. Thus, we are locally writing the distribution as

$$\mathcal{D} = \{(r, s, \dot{r}, \dot{s}) \in TQ \mid \dot{s} + A_\alpha^a \dot{r}^\alpha = 0\}.$$

The equations of motion, (3.1) may be rewritten by noting that the allowed variations $\delta q^i = (\delta r^\alpha, \delta s^a)$ satisfy $\delta s^a + A_\alpha^a \delta r^\alpha = 0$. Substitution into (3.1) gives

$$\left(\frac{d}{dt} \frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial r^\alpha} \right) = A_\alpha^a \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a} \right). \quad (3.2)$$

Equation (3.2) combined with the constraint equations

$$\dot{s}^a = -A_\alpha^a \dot{r}^\alpha \quad (3.3)$$

gives a complete description of the equations of motion of the system; this procedure may be viewed as one way of eliminating the Lagrange multipliers. Using this notation, one finds that $\lambda = \lambda_a \omega^a$, where

$$\lambda_a = \frac{d}{dt} \frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a}.$$

Equations (3.2) can be written in the following way:

$$\frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A_\alpha^a \frac{\partial L_c}{\partial s^a} = - \frac{\partial L}{\partial \dot{s}^b} B_{\alpha\beta}^b \dot{r}^\beta, \quad (3.4)$$

where

$$L_c(r^\alpha, s^a, \dot{r}^\alpha) = L(r^\alpha, s^a, \dot{r}^\alpha, -A_\alpha^a(r, s) \dot{r}^\alpha).$$

is the coordinate expression of the constrained Lagrangian defined by $L_c = L|_{\mathcal{D}}$ and where

$$B_{\alpha\beta}^b = \left(\frac{\partial A_\alpha^b}{\partial r^\beta} - \frac{\partial A_\beta^b}{\partial r^\alpha} + A_\alpha^a \frac{\partial A_\beta^b}{\partial s^a} - A_\beta^a \frac{\partial A_\alpha^b}{\partial s^a} \right). \quad (3.5)$$

Letting $\mathbf{d}\omega^b$ be the exterior derivative of ω^b , a computation shows that

$$\mathbf{d}\omega^b(\dot{q}, \cdot) = B_{\alpha\beta}^b \dot{r}^\alpha dr^\beta$$

and hence the equations of motion have the form

$$-\delta L_c = \left(\frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A_\alpha^a \frac{\partial L_c}{\partial s^a} \right) \delta r^\alpha = - \frac{\partial L}{\partial \dot{s}^b} d\omega^b(\dot{q}, \delta r).$$

This form of the equations isolates the effects of the constraints, and shows, in particular, that in the case where the constraints are integrable (*i.e.*, $\mathbf{d}\omega = 0$), the equations of motion are obtained by substituting the constraints into the Lagrangian and then setting the variation of L_c to zero. However in the non-integrable case the constraints generate extra (curvature) terms, which must be taken into account.

The above coordinate results can be put into an interesting and useful intrinsic geometric framework. The intrinsically given information is the distribution and the Lagrangian. *Assume temporarily* that there is a bundle structure $\pi_{Q,R} : Q \rightarrow R$ for our space Q , where R is the *base* manifold and $\pi_{Q,R}$ is a submersion and the kernel of $T_q\pi_{Q,R}$ at any point $q \in Q$ is called the *vertical space* V_q . One can always do this locally. An **Ehresmann connection** A is a vertical valued one form on Q such that

1. $A_q : T_q Q \rightarrow V_q$ is a linear map and
2. A is a projection: $A(v_q) = v_q$ for all $v_q \in V_q$.

Hence, $T_q Q = V_q \oplus H_q$ where $H_q = \ker A_q$ is the **horizontal space at q** , sometimes denoted hor_q . Thus, an Ehresmann connection gives us a way to split the tangent space to Q at each point into a horizontal and vertical part.

If the Ehresmann connection is chosen in such a way that the given constraint distribution \mathcal{D} is the horizontal space of the connection; that is, $H_q = \mathcal{D}_q$, then in the bundle coordinates $q^i = (r^\alpha, s^a)$, the map $\pi_{Q,R}$ is just projection onto the factor r and the connection A can be represented locally by a vector valued differential form ω^a :

$$A = \omega^a \frac{\partial}{\partial s^a}, \quad \omega^a(q) = ds^a + A_\alpha^a(r, s) dr^\alpha,$$

and the horizontal projection is the map

$$(\dot{r}^\alpha, \dot{s}^a) \mapsto (\dot{r}^\alpha, -A_\alpha^a(r, s)\dot{r}^\alpha).$$

The **curvature** of an Ehresmann connection A is the vertical valued two form defined by its action on two vector fields X and Y on Q as

$$B(X, Y) = -A([\text{hor } X, \text{hor } Y])$$

where the bracket on the right hand side is the Jacobi-Lie bracket of vector fields obtained by extending the stated vectors to vector fields. This definition shows the sense in which the curvature measures the failure of the constraint distribution to be integrable.

In coordinates, one can evaluate the curvature B of the connection A by the following formula:

$$B(X, Y) = \mathbf{d}\omega^a(\text{hor } X, \text{hor } Y) \frac{\partial}{\partial s^a},$$

so that the local expression for curvature is given by

$$B(X, Y)^a = B_{\alpha\beta}^a X^\alpha Y^\beta$$

where the coefficients $B_{\alpha\beta}^a$ are given by (3.5).

The Lagrange d'Alembert equations may be written intrinsically as

$$\delta L_c = \langle \mathbb{F}L, B(\dot{q}, \delta q) \rangle,$$

in which δq is a horizontal variation (*i.e.*, it takes values in the horizontal space) and B is the curvature regarded as a vertical valued two form, in addition to the constraint equations $A(q) \cdot \dot{q} = 0$. Here $\langle \cdot, \cdot \rangle$ denotes the pairing between a vector and a dual vector and

$$\delta L_c = \left\langle \delta r^\alpha, \frac{\partial L_c}{\partial r^\alpha} - \frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^\alpha} - A_\alpha^a \frac{\partial L_c}{\partial s^a} \right\rangle.$$

When there is a symmetry group G present, there is a natural bundle one can work with and put a connection on, namely the bundle $Q \rightarrow Q/G$. In the generality of the preceding discussion, one can get away with just the distribution itself and can introduce the corresponding Ehresmann connection locally. In fact, the bundle structure $Q \rightarrow R$ is really a “red herring”. The notion of curvature as a $T_q Q / \mathcal{D}_q$ valued form makes good sense and is given locally by the same expressions as above. However, keeping in mind that we eventually want to deal with symmetries and in that case there is a natural bundle, the Ehresmann assumption is nevertheless a reasonable bridge to the more interesting case with symmetries.

More on the Euler-Poincaré equations. There is a way to carry out reduction for a nonholonomic system. We shall focus on the Lagrangian side; for the equivalent Poisson picture, see Koon and Marsden [1998b,c].

The idea is to pass the Lagrange d'Alembert principle to \mathcal{D}/G ; in a way that is similar to the reduction of Hamilton's principle (e.g., giving the *Routhian*, etc.), as mentioned in the first lecture. Recall that a simple example of Lagrangian reduction is the *free rigid body*; the Euler equations, namely $I\dot{\Omega} = I\Omega \times \Omega$, are *not variational*, but they satisfy a Lagrange d'Alembert type of principle (*i.e.*, there are constraints on the allowed variations) that is obtained by reducing Hamilton's principle from $\text{SO}(3)$.

To enhance the discussion in the first lecture on the Euler-Poincaré equations, we now present an extension of them, following Holm, Marsden and Ratiu [1998a]. In fact, this provides a Lagrangian version of the semidirect product reduction theory discussed in the last lecture. Following this, we return to the question of how these procedures work for nonholonomic systems.

The basic ingredients we start with are as follows. Assume there is a *left* representation of Lie group G on the vector space V ; then G acts on $TG \times V^*$ as well. Assume $L : TG \times V^* \rightarrow \mathbb{R}$ is left G -invariant. For $a_0 \in V^*$, define $L_{a_0} : TG \rightarrow \mathbb{R}$ by

$$L_{a_0}(v_g) = L(a_0, v_g),$$

and define $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$ by

$$l(g^{-1}v_g, g^{-1}a) = L(v_g, a).$$

For a curve $g(t) \in G$, let

$$\xi(t) := g(t)^{-1}\dot{g}(t)$$

and define the curve $a(t)$ as the unique solution of the equation

$$\dot{a}(t) = -\xi(t)a(t)$$

with initial condition $a(0) = a_0$; *i.e.*, $a(t) = g(t)^{-1}a_0$.

Theorem 3.1 *The following assertions are equivalent:*

1. With a_0 fixed, the standard Hamilton principle holds:

$$\delta \int_{t_1}^{t_2} L_{a_0}(g(t), \dot{g}(t)) dt = 0$$

for variations with fixed endpoints.

2. The curve $g(t)$ satisfies the standard Euler-Lagrange equations for L_{a_0} .
3. The Lagrange d'Alembert-type principle

$$\delta \int_{t_1}^{t_2} l(\xi(t), a(t)) dt = 0$$

holds on \mathfrak{g} , using variations of ξ and a of the form

$$\delta \xi = \dot{\eta} + [\xi, \eta], \quad \delta a = -\eta a,$$

where $\eta(t)$ is a curve in \mathfrak{g} vanishing at the endpoints.

4. The **Euler-Poincaré** equations hold on $\mathfrak{g} \times V^*$

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a$$

where $\rho_v : \mathfrak{g} \rightarrow V$ is given by $\xi \mapsto \xi v$, the infinitesimal action, $\rho_v^* : V^* \rightarrow \mathfrak{g}^*$ is its dual, and $v \diamond a = \rho_v^* a$.

This is proved by a straightforward adaptation of the proof of Euler-Poincaré reduction that was presented in the first lecture. The main extra feature is to keep track of the “constraint” $a_0 = \text{constant}$. Since $a_0 = ga$, we get $0 = (\delta g)a + g\delta a$, i.e., $\delta a = -(g^{-1}\delta g)a = -\eta a$.

Nonholonomic reduced equations. Reduction procedures were applied to the Lagrange d’Alembert principle in Bloch, Krishnaprasad, Marsden and Murray [1996]. The *form of the resulting reduced equations* is

$$\begin{aligned} g^{-1}\dot{g} &= -A(r)\dot{r} + B(r)p, \\ \dot{p} &= \dot{r}^T \alpha(r)\dot{r} + \dot{r}^T \beta(r)p + p^T \gamma(r)p \\ M(r)\ddot{r} &= -C(r, \dot{r}) + N(r, \dot{r}, p) + \tau \end{aligned}$$

The *first equation* describes the *motion in the group variables* as the flow of a left-invariant vector field on G determined by the *internal shape* r , the *internal velocity* \dot{r} , as well as the *nonholonomic momentum* p , which is, roughly speaking, the component of momentum in the symmetry directions compatible with the constraints. The *second equation* is the *momentum equation*. Notice that the momentum equation has terms that are

- quadratic in \dot{r} ,
- linear in \dot{r} and p and
- quadratic in p .

The coefficients $\beta(r)$ define a *connection* and this term is called the *transport part* of the momentum equation. The *curvature* of this connection plays an important role in stability theory. The *third equation* describes the *motion in the shape variables* r . The term $M(r)$ is the *mass matrix* of the system, C is the *Coriolis term* and τ represent internal control forces. This framework has proven to be useful for *controllability, gait selection, and locomotion* of systems like the snakeboard. We will come back to this below.

A nice example illustrating many features of nonholonomic systems is the bicycle (see Figure 3.4). This example (under drastically simplified modeling assumptions) is studied in Koon and Marsden [1998a]. We also mention that Koon and Marsden [1998c] study the above reduced Lagrange d’Alembert equations from the Poisson point of view.

The snakeboard. The snakeboard is an interesting example of a system in which there is a nontrivial interaction between the forces of constraint and the momentum laws that arise due to symmetries. As such, it has played an important role in understanding the many subtleties of nonholonomic systems.

The *snakeboard* is similar to a skateboard, with one exception: the front and back pairs of wheels can be rotated independently about their vertical axes. Thus, a rider stands with one foot above each of the wheel bases, and can couple twisting motions of the torso with turning of the feet. The most interesting aspect of this motion is that the rider can begin to move forward without ever having to kick off the ground or pedal, only needing to

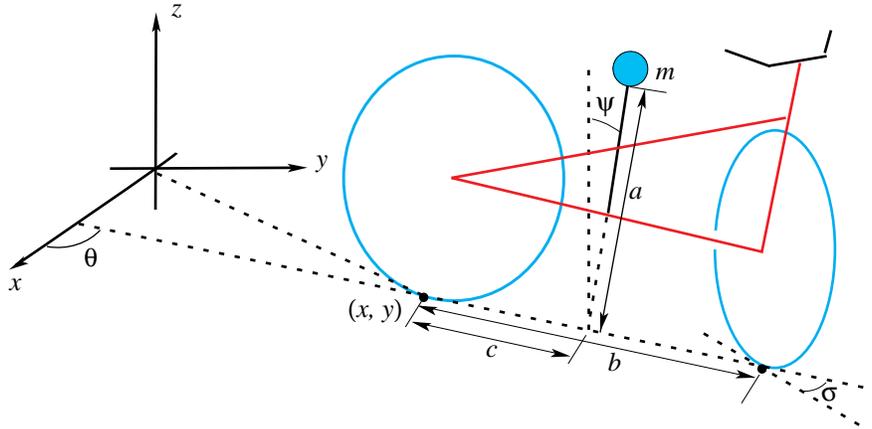


Figure 3.4: The mechanics of a bicycle (from Koon and Marsden [1998a]).

coordinate the twisting gyrations with the rotation of the wheel axles. The resultant path that is traced out is similar to the serpentine motion of a snake, thus the name *snakeboard*.

In Figure 3.5 we show a model of the snakeboard, in which the twisting of the torso has been replaced by a rotating inertia wheel.

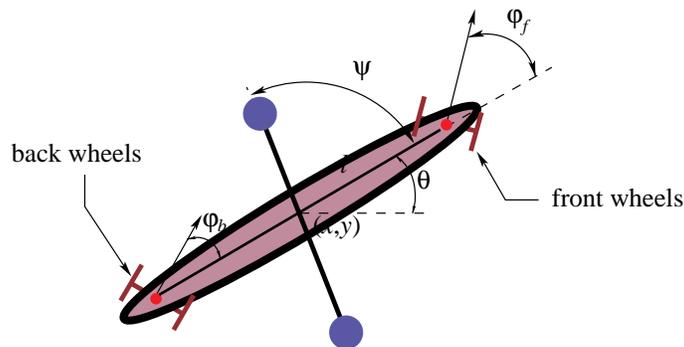


Figure 3.5: A simplified model of the *snakeboard*.

This model has been studied in detail by Lewis, Ostrowski, Murray and Burdick [1994] and Bloch, Krishnaprasad, Marsden and Murray [1996]. From the theoretical point of view, the feature of the snakeboard that sets it apart from examples like the planar skater and the falling cat is that even though it has the symmetry group of rotations and translations of the plane, the *linear and angular momentum is not conserved*.

For the planar skater, no matter what motions the arms of the device make, the values of the linear and angular momentum cannot be altered. Thus, while it is possible to change the orientation of the planar skater, once the internal shape motions stop, the orientation changes also stop. This is not true for the snakeboard—one has the ability to build up momentum, which can be traced to the presence of forces of constraint. Thus, one might suspect that one should abandon the ideas of linear and angular momentum for the snakeboard. However, a deeper inspection shows that this is not the case. In fact, one finds that there is a particular “component” of the angular momentum, namely (a multiple of) that

about the point O shown in Figure 3.6 that satisfies a special equation. The precise relationship for the snakeboard, as comes about from a little detective work on the definitions, is $p = (\text{angular momentum}) \times \sin(2\phi)$.

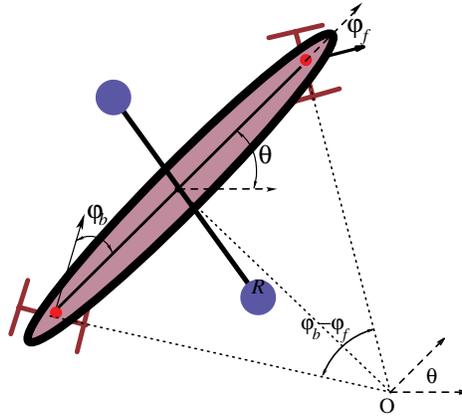


Figure 3.6: The angular momentum about the point O plays an important role in the analysis of the snakeboard.

If we call this quantity p the *nonholonomic momentum*, one finds that due to the translational and rotational invariance of the whole system, there is a *nonholonomic momentum equation* governing the evolution of p , which has the form given earlier. Assuming $\phi_f = -\phi_b = \phi$ for simplicity, the momentum equation is

$$\dot{p} = 2J_0(\cos^2 \phi)\dot{\phi}\dot{\psi} - (\tan \phi)p\dot{\phi},$$

where ϕ and ψ represent the internal variables of the system and J_0 is the rotor inertia. An important point to recognize is that this equation does not depend on the rotational and translational position of the system, *i.e.*, there is no explicit g dependence *i.e.*, no x, y or θ dependence, which parameterize overall translations and rotations of the system. Thus, if one has a given internal motion, this equation can be solved for p and from it, the attitude and position of the snakeboard calculated by means of another integration using the reconstruction equation for $g^{-1}\dot{g}$. This strategy thus parallels that used to study the falling cat and the planar skater.

Locomotion and gaits. Material in this section is adapted from Marsden and Ostrowski [1998]. The snakeboard moves by coupling periodic motions of the rotor and wheel axles. Similarly, the roller racer moves using coupled internal motions. In each case, an understanding of the geometry behind this sort of mechanism has proven to be quite useful.

Special periodic internal motions that generate specific types of locomotion are examples of *gaits*, which more generally can be thought of as cyclic patterns of internal shape changes which result in a net displacement. Of course the term is adapted from the animal world where similar things happen, but in fact are more subtle than what we are discussing here due, in part to the fact that in those examples there are periodic changes in the phase space as the legs of the animal come into periodic contact with the ground.

Different gaits correspond to different cyclic input patterns. For example, the snakeboard possesses at least three primary gait patterns, shown in Figures 3.7–3.9. The first is the *drive gait*, shown in Figure 3.7. This figure shows the position of the snakeboard's center

of mass versus time for the case in which the rotor and wheel axles oscillate with the same frequency (which is called the “drive gait”). This gait closely resembles the motion followed by riders of the *snakeboard* when they begin moving from a resting position.

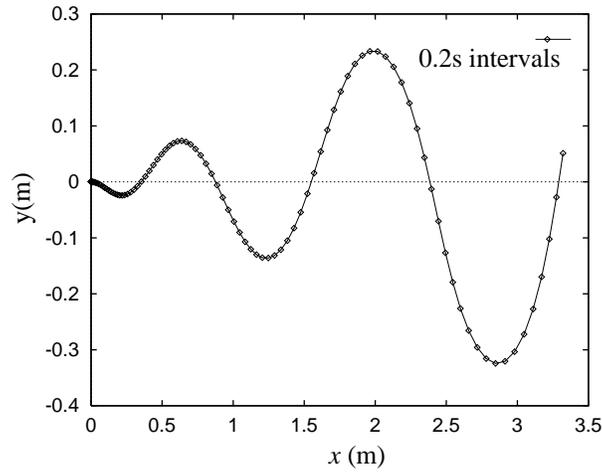


Figure 3.7: Center of mass position for the drive gait.

Fig. 3.8 shows a second possible gait, the *rotate gait*, which has not yet been discovered by *snakeboard* riders, but which is easily demonstrated by a robotic version that has been built. In this case, the rotor oscillates at twice the frequency of the axles (the “rotate gait,” as the robot essentially rotates in place).

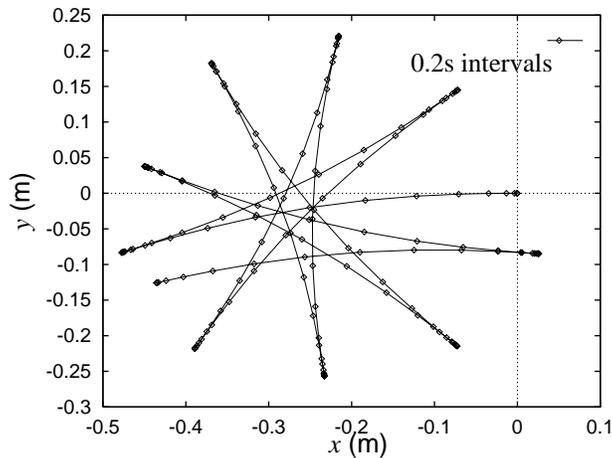


Figure 3.8: Position of the center of mass for the “rotate gait”.

Finally, in Figure 3.9 we show the path followed when generating the third direction of motion (having already generated forward translation and rotation). This gait, the *parallel parking gait*, is the most complicated to perform, requiring that the rotor oscillate three times for every two oscillations of the axles.

We do not yet have a complete geometric understanding of the notion of gaits. In general, the net displacement of the mechanism that arises from periodic inputs has, as an

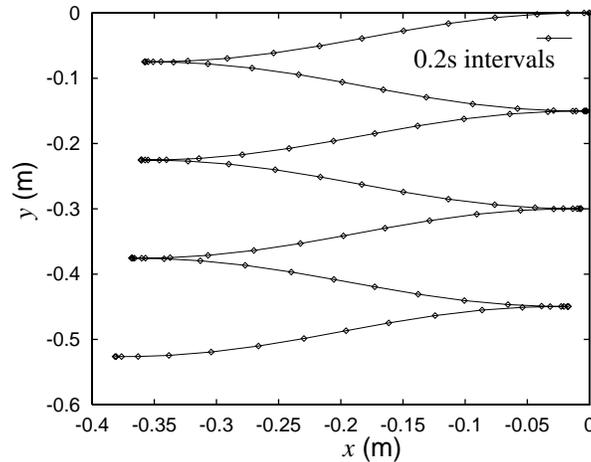


Figure 3.9: Position of the center of mass for the “parallel parking gait”.

important ingredient, the *geometric phase*, or holonomy. The geometric phase is that part of the motion described by the local form of the mechanical connection, \mathcal{A} . In the case of the snakeboard, the net displacement is a non-trivial combination of the geometric phase and the dynamic phase. Understanding the increased complexity of the relationship between geometric and dynamic phases for this class of systems is the subject of current research. However, it has been found by direct calculation that different gaits can be associated with the derivatives of the connection. When the momentum terms are not present, *i.e.*, when $p \equiv 0$ (the case of kinematic constraints), all motions are generated through this mechanism. In the presence of momentum, however, there is a tight coupling between the geometry of the connection and the effect of the inertia and momentum. This suggests a very strong relationship between the geometric and dynamic phases as encoded by the connection, and the generation of gaits.

Stability, controllability, and optimal control. Control theory adds to the study of dynamical systems the idea that in many instances, one can directly intervene in the dynamics rather than passively watching. For example, while Newton’s equations govern the dynamics of a satellite, we can intervene in these dynamics by controlling onboard gyroscopes, thrusters, or rotors. Quite often, control engineers are tempted to overwhelm the intrinsic dynamics of a system with the controls. However, in many circumstances (fluid control is an example—see, for example, the discussion in Bloch and Marsden [1989]) one needs to work with the intrinsic dynamics and make use of its structure.

Two of the basic notions in control theory involve steering and stabilizability. Steering has as its objective, the design of control inputs that guide the system from an initial position to a desired final position, perhaps following a predefined path. One imagines manipulating the control to achieve this, much the way one steers a car so that the desired final state is achieved. Already this type of question has received much attention and many important and basic questions have been answered. For example, two of the main themes that have developed are the Lie algebraic techniques based on brackets of vector fields (in driving a car, you can repeatedly make two alternating steering motions to produce a motion in a third direction as one does in parallel parking) and the second based on the application of exterior differential systems (a subject invented by Elie Cartan in the mid 1920s whose power is only now being significantly tapped in control theory). The work of Tilbury, Murray, and Sastry

[1993], and Walsh and Bushnell [1993], and Leonard and Krishnaprasad [1995] typify some of the modern applications of these ideas.

We shall go into the analysis and geometry of some optimal control problems, and in particular, the falling cat problem in the next lecture.

Lecture 4

Optimal Control and Stabilization of Balance Systems

In this lecture we consider two problems from control theory. The first is an optimal control problem, leading up to the falling cat theorem while the second is a stabilization problem.

Optimal control for the Heisenberg system. We first illustrate some basic techniques in a simple example called the Heisenberg system. This example, due to Brockett [1981], is a prototype for the falling cat problem and a variety of optimal steering problems.

The *Heisenberg system* is the following control system in \mathbb{R}^3 :

$$\begin{aligned}\dot{x} &= u_1 \\ \dot{y} &= u_2 \\ \dot{z} &= xu_2 - yu_1,\end{aligned}$$

where u_1 and u_2 are *control inputs*. The system may be written as

$$\dot{q} = u_1 g_1 + u_2 g_2$$

where $q = (x, y, z)^T$, $g_1 = (1, 0, y)^T$ and $g_2 = (0, 1, -x)^T$. Note that g_1 and g_2 are a set of independent vector fields satisfying the constraint

$$\dot{z} = x\dot{y} - y\dot{x}.$$

One verifies that the Jacobi-Lie bracket of the vector fields g_1 and g_2 is

$$[g_1, g_2] = 2g_3$$

where $g_3 = (0, 0, 1)$. In fact, the three vector fields g_1, g_2, g_3 span all of \mathbb{R}^3 and, as a Lie algebra, is just the Heisenberg algebra for the basic operators (up to scale factors) q, p and the identity, from quantum mechanics.

By general controllability theorems (Chow's theorem) that are closely related to the Frobenius theorem, one knows that one can, with a suitable choice of controls u_1 and u_2 , steer trajectories between any two points in \mathbb{R}^3 . In particular, we are interested in the following optimal steering problem (see Figure 4.1):

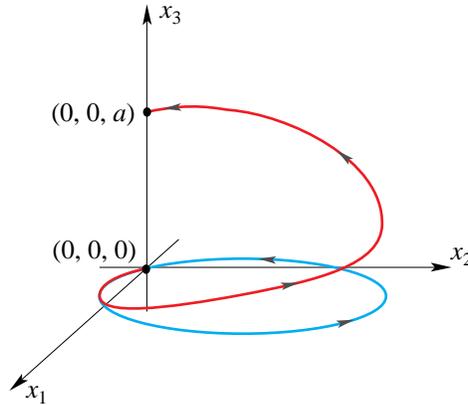


Figure 4.1: An optimal steering problem.

Optimal steering problem. Given numbers $a > 0$ and $T > 0$, find time dependent controls u_1, u_2 that steer the trajectory starting at $(0, 0, 0)$ at time $t = 0$ to the point $(0, 0, a)$ after time $T > 0$ and that, amongst all such controls, minimizes

$$\frac{1}{2} \int_0^T (u_1^2 + u_2^2) dt.$$

Solution of the optimal steering problem. An equivalent formulation is the following: minimize the integral

$$\frac{1}{2} \int_0^T (\dot{x}^2 + \dot{y}^2) dt$$

amongst all curves $x(t)$ joining $x(0) = (0, 0, 0)$ to $x(T) = (0, 0, a)$ that satisfy the constraint

$$\dot{z} = y\dot{x} - x\dot{y}.$$

The calculus of variations analogue of the Lagrange multiplier theorem states that any solution must satisfy the Euler-Lagrange equations for the Lagrangian with a Lagrange multiplier inserted:

$$L(x, \dot{x}, y, \dot{y}, z, \dot{z}, \lambda, \dot{\lambda}) = \frac{1}{2} (\dot{x}^2 + \dot{y}^2) + \lambda (\dot{z} - y\dot{x} + x\dot{y})$$

The corresponding Euler-Lagrange equations are given by

$$\begin{aligned} \ddot{x} - 2\lambda\dot{y} &= 0 \\ \ddot{y} + 2\lambda\dot{x} &= 0 \\ \dot{\lambda} &= 0 \\ \dot{z} - y\dot{x} + x\dot{y} &= 0. \end{aligned}$$

From the third equation, λ is a constant, and the first two equations state that *the particle* $(x(t), y(t))$ moves in the plane in a constant magnetic field (pointing in the z direction, with charge proportional to the constant λ).

1. Note that this optimal steering problem gives rise to an interesting mechanical system, a particle in a magnetic field.
2. Since particles in constant magnetic fields move in circles with constant speed, they have a sinusoidal time dependence, and hence so do the controls. This has led to the “steering by sinusoids” approach in many nonholonomic steering problems (see for example Murray and Sastry [1993]).

The equations for x and y are linear first order equations in the velocities and are readily solved:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} = \begin{bmatrix} \cos(2\lambda t) & \sin(2\lambda t) \\ -\sin(2\lambda t) & \cos(2\lambda t) \end{bmatrix} \begin{bmatrix} \dot{x}(0) \\ \dot{y}(0) \end{bmatrix}.$$

Integrating once more and using the initial conditions $x(0) = 0, y(0) = 0$ gives:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \frac{1}{2\lambda} \begin{bmatrix} \cos(2\lambda t) - 1 & \sin(2\lambda t) \\ -\sin(2\lambda t) & \cos(2\lambda t) - 1 \end{bmatrix} \begin{bmatrix} -\dot{y}(0) \\ \dot{x}(0) \end{bmatrix}.$$

The other boundary condition $x(T) = 0, y(T) = 0$ gives

$$\lambda = \frac{n\pi}{T}.$$

Using this information, we find z by integration: from $\dot{z} = x\dot{y} - y\dot{x}$ and the preceding expressions, we get

$$\dot{z}(t) = \frac{1}{2\lambda} [-\dot{x}(0)^2 - \dot{y}(0)^2 + \cos(2\lambda t)(\dot{x}(0)^2 + \dot{y}(0)^2)].$$

Integration from 0 to T and using $z(0) = 0$ gives

$$z(T) = \frac{T}{2\lambda} [-\dot{x}(0)^2 - \dot{y}(0)^2].$$

Thus, to achieve the boundary condition $z(T) = a$, one chooses

$$\dot{x}(0)^2 + \dot{y}(0)^2 = -\frac{2\pi na}{T^2}.$$

One also finds that

$$\begin{aligned} \frac{1}{2} \int_0^T [\dot{x}(t)^2 + \dot{y}(t)^2] dt &= \frac{1}{2} \int_0^T [\dot{x}(0)^2 + \dot{y}(0)^2] dt \\ &= \frac{T}{2} [\dot{x}(0)^2 + \dot{y}(0)^2] \\ &= -\frac{\pi na}{T}, \end{aligned}$$

so that the minimum is achieved when $n = -1$.

Summary. The *solution of the optimal control problem* is given by choosing initial conditions such that $\dot{x}(0)^2 + \dot{y}(0)^2 = 2\pi a/T^2$ and with the trajectory in the xy plane given by the circle

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \frac{1}{2\lambda} \begin{bmatrix} \cos(2\pi t/T) - 1 & -\sin(2\pi t/T) \\ \sin(2\pi t/T) & \cos(2\pi t/T) - 1 \end{bmatrix} \begin{bmatrix} -\dot{y}(0) \\ \dot{x}(0) \end{bmatrix}$$

and with z given by

$$z(t) = \frac{ta}{T} - ta^2 \sin\left(\frac{2\pi t}{T}\right).$$

Notice that any such solution can be rotated about the z axis to obtain another one.

Falling cat problem. This problem is an abstraction of the problem of how a falling cat should optimally (in some sense) move its body parts so that it achieves a 180° reorientation during its fall.

We begin with a Riemannian manifold Q (the configuration space of the problem) with a free and proper isometric action of a Lie group G on Q (the group $SO(3)$ for the falling cat). Let \mathcal{A} denote the mechanical connection; that is, it is the principal connection whose horizontal space is the metric orthogonal to the group orbits. The quotient space $Q/G = X$, the shape space, inherits a Riemannian metric from that on Q . Given a curve $c(t)$ in Q , we shall denote the corresponding curve in the base space X by $r(t)$.

The optimal control problem under consideration is as follows:

Isoholonomic problem (falling cat problem). *Fixing two points $q_1, q_2 \in Q$, among all curves $q(t) \in Q$, $0 \leq t \leq 1$ such that $q(0) = q_0, q(1) = q_1$ and $\dot{q}(t) \in \text{hor}_{q(t)}$ (horizontal with respect to the mechanical connection \mathcal{A}), find the curve or curves $q(t)$ such that the energy of the base space curve, namely,*

$$\frac{1}{2} \int_0^1 \|\dot{r}\|^2 dt,$$

is minimized.

Theorem 4.1 (Montgomery [1984, 1990, 1991a]). *If $q(t)$ is a (regular) optimal trajectory for the isoholonomic problem, then there exists a curve $\lambda(t) \in \mathfrak{g}^*$ such that the reduced curve $r(t)$ in $X = Q/G$ together with $\lambda(t)$ satisfies **Wong's equations**:*

$$\begin{aligned} \dot{p}_\alpha &= -\lambda_a \mathcal{B}_{\alpha\beta}^a \dot{r}^\beta - \frac{1}{2} \frac{\partial g^{\beta\gamma}}{\partial r^\alpha} p_\beta p_\gamma \\ \dot{\lambda}_b &= -\lambda_a C_{db}^a \mathcal{A}_\alpha^d \dot{r}^\alpha \end{aligned}$$

where $g_{\alpha\beta}$ is the local representation of the metric on the base space X ; that is

$$\frac{1}{2} \|\dot{r}\|^2 = \frac{1}{2} g_{\alpha\beta} \dot{r}^\alpha \dot{r}^\beta,$$

$g^{\beta\gamma}$ is the inverse of the matrix $g_{\alpha\beta}$, p_α is defined by

$$p_\alpha = \frac{\partial l}{\partial \dot{r}^\alpha} = g_{\alpha\beta} \dot{r}^\beta,$$

and where we write the components of \mathcal{A} as \mathcal{A}_α^b and similarly for its curvature \mathcal{B} .

Proof. As with the Heisenberg system, by general principles in the calculus of variations, given an optimal solution $q(t)$, there is a Lagrange multiplier $\lambda(t)$ such that the new action function defined on the space of curves with fixed endpoints by

$$\mathfrak{S}[q(\cdot)] = \int_0^1 \left[\frac{1}{2} \|\dot{r}(t)\|^2 + \langle \lambda(t), \mathcal{A}\dot{q}(t) \rangle \right] dt$$

has a critical point at this curve. Using the integrand as a Lagrangian, identifying $\Omega = \mathcal{A}\dot{q}$ and applying the reduced Euler-Lagrange equations from Lecture 1 to the reduced Lagrangian

$$l(r, \dot{r}, \Omega) = \frac{1}{2} \|\dot{r}\|^2 + \langle \lambda, \Omega \rangle$$

then gives Wong's equations by the following simple calculations:

$$\frac{\partial l}{\partial \dot{r}^\alpha} = g_{\alpha\beta} \dot{r}^\beta; \quad \frac{\partial l}{\partial r^\alpha} = \frac{1}{2} \frac{\partial g^{\beta\gamma}}{\partial r^\alpha} \dot{r}^\beta \dot{r}^\gamma; \quad \frac{\partial l}{\partial \Omega^a} = \lambda_a.$$

The constraints are $\Omega = 0$ and so the reduced Euler-Lagrange equations become

$$\begin{aligned} \frac{d}{dt} \frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} &= -\lambda_a (\mathcal{B}_{\alpha\beta}^a \dot{r}^\beta) \\ \frac{d}{dt} \lambda_b &= -\lambda_a (\mathcal{E}_{\alpha b}^a \dot{r}^\alpha) = -\lambda_a C_{db}^a \mathcal{A}_\alpha^d \dot{r}^\alpha. \end{aligned}$$

But

$$\begin{aligned} \frac{d}{dt} \frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} &= \dot{p}_\alpha - \frac{1}{2} \frac{\partial g_{\beta\gamma}}{\partial r^\alpha} \dot{r}^\beta \dot{r}^\gamma \\ &= \dot{p}_\alpha + \frac{1}{2} \frac{\partial g^{\kappa\sigma}}{\partial r^\alpha} g_{\kappa\beta} g_{\sigma\gamma} \dot{r}^\beta \dot{r}^\gamma \\ &= \dot{p}_\alpha + \frac{1}{2} \frac{\partial g^{\beta\gamma}}{\partial r^\alpha} p_\beta p_\gamma, \end{aligned}$$

and so we have the desired equations. ■

Remark. There is a rich literature on Wong's equations and it was an important ingredient in the development of reduction theory. Some references are Sternberg [1977], Guillemin and Sternberg [1978], Weinstein [1978], Montgomery, Marsden and Ratiu [1984], Montgomery [1984], Koon and Marsden [1997] and Cendra, Holm, Marsden, and Ratiu [1998].

Nonholonomic optimal control. Using a synthesis of the techniques used above for the Heisenberg system and the falling cat problem, Koon and Marsden [1997] generalized these problems to the nonholonomic case. In addition, these methods allow one to treat the falling cat problem even in the case that the angular momentum is not zero.

In this process the momentum equation plays the role of the constraint. It is inserted as a first order differential constraint on the nonholonomic momentum.

Stabilizing the inverted pendulum. We next illustrate the stabilization method of Bloch, Leonard and Marsden [1997] for a *pendulum on a cart*. The *inverted spherical pendulum* is a little more complicated, but the methods also work in this case. They also work in many other cases, such as satellites with internal rotors, underwater vehicles with rotors, etc. (See Bloch, Leonard and Marsden [1998]). Our methods are similar in spirit to energy methods that are proving very effective (see, e.g., Åström and Furuta [1997]).

Our idea for stabilization is to use the mechanical structure and to *create an energy extrema*. Dissipation (real or control) can then *convert stability to asymptotic stability*. The main complication for more complex systems such as the inverted spherical pendulum, the satellite and the underwater vehicle are *gyroscopic forces*. Reduction techniques and in particular, magnetic terms that arise from reduction are perfect tools for investigating such problems.

Lagrangian for the cart-pendulum system. Let s denote the position of the cart on the s -axis and let θ denote the angle of the pendulum from the upright vertical, as in Figure 4.2.

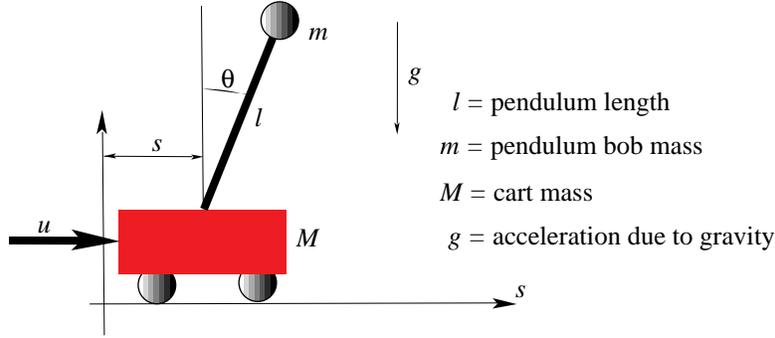


Figure 4.2: The pendulum on a cart system.

The *configuration space* is $Q = \mathbb{R} \times S^1$ and the *velocity phase space*, TQ has coordinates $(s, \theta, \dot{s}, \dot{\theta})$. The velocity of the cart relative to the lab frame is \dot{s} and the velocity of the pendulum is the vector $(\dot{s} + l \cos \theta \dot{\theta}, -l \sin \theta \dot{\theta})$. Thus, the *system kinetic energy* is

$$K((s, \theta, \dot{s}, \dot{\theta})) = \frac{1}{2}(\dot{s}, \dot{\theta}) \begin{pmatrix} M + m & ml \cos \theta \\ ml \cos \theta & ml^2 \end{pmatrix} \begin{pmatrix} \dot{s} \\ \dot{\theta} \end{pmatrix}.$$

and so the Lagrangian is

$$L(s, \theta, \dot{s}, \dot{\theta}) = K(s, \theta, \dot{s}, \dot{\theta}) - V(\theta),$$

where the potential energy is $V = mgl \cos \theta$.

The *symmetry group* is that of translation in the variable s , so $G = \mathbb{R}$. We do not destroy this symmetry when doing stabilization in θ .

Controlled Cart. Write the above Lagrangian as

$$L(s, \theta, \dot{s}, \dot{\theta}) = \frac{1}{2} \left(\alpha \dot{\theta}^2 + 2\beta \cos \theta \dot{s} \dot{\theta} + \gamma \dot{s}^2 \right) + D \cos \theta,$$

where $\gamma = M + m$, $\beta = ml$, $\alpha = ml^2$ and $D = -mgl$. Positive definiteness of the mass matrix (the Riemannian metric) corresponds to $\alpha\gamma - \beta^2 > 0$. The *momentum conjugate* to s is $p_s = \gamma \dot{s} + \beta \cos \theta \dot{\theta}$. The *relative equilibrium* $\theta = 0, \dot{\theta} = 0$ is unstable since $D < 0$. This upright state is what we wish to stabilize. The *equations of motion* of the cart-pendulum system subject to a control force u acting on the cart (and no direct forces on the pendulum) are the *controlled cart equations*:

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{s}} &= u \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} &= 0 \end{aligned}$$

Controlled Lagrangian. The *controlled Lagrangian* is defined by modifying the kinetic energy only (potentials are reserved for tracking). Let σ be a real scalar and introduce a one form $\tau = k(\theta)d\theta$. Define

$$L_{\tau, \sigma} = \frac{1}{2} \left[\alpha \dot{\theta}^2 + 2\beta \cos \theta (\dot{s} + k(\theta)\dot{\theta})\dot{\theta} + \gamma (\dot{s} + k(\theta)\dot{\theta})^2 \right] + \frac{\sigma}{2} \gamma [k(\theta)]^2 \dot{\theta}^2 + D \cos \theta.$$

Note that the variable s is “shifted” by the one form τ and a term quadratic in $\dot{\theta}$ is added. There is a *general construction* of new Lagrangians of this type based on changes of connections (velocity shifts) and using *Kaluza-Klein ideas*.

The variable s is still *cyclic*. The associated conservation law for the controlled Lagrangian is

$$\frac{d}{dt}p_s = u := -\gamma \frac{d}{dt}[k(\theta)\dot{\theta}]$$

We strive to identify the term on the right hand side with the **control force** (the force exerted on the cart). A (relatively miraculous) computation shows that *the θ equation for the controlled Lagrangian matches with the θ equation for the controlled cart if we choose $k(\theta) = \kappa \cos \theta$ for a real number κ and if we choose $\sigma = -\beta/(\gamma\kappa)$. The resulting control law is*

$$u = \gamma\kappa \left(\sin \theta \dot{\theta}^2 + \cos \theta f(\theta, \dot{\theta}) \right),$$

where

$$f(\theta, \dot{\theta}) = \frac{D \sin \theta + \dot{\theta}^2 \left(\frac{\beta^2}{\gamma} + \beta\kappa \right) \cos \theta \sin \theta}{\alpha - \left(\frac{\beta^2}{\gamma} - \beta\kappa \right) \cos^2 \theta}.$$

In linear approximation it is a proportional controller, $u = \text{constant} \times \theta$.

Stabilization. The θ dynamics is stabilized if the energy has an extremum at the equilibrium (in this case a maximum); this leads to a condition on the “nonlinear gain” κ :

$$\kappa > \frac{\alpha\gamma - \beta^2}{\beta\gamma} > 0.$$

Having a maximum and not a minimum is not a problem even if dissipation is present—simulate negative dissipation by the controller; one then gets *asymptotic stability* instead of *Liapunov stability*.

Summary for the inverted pendulum. *We get a stabilizing feedback control law (an expression for the control force needed as a function of the state of the pendulum) provided κ is chosen to satisfy the preceding inequality. Stability is determined by energy considerations. This procedure allows one to **discover** the stabilizing control law as long as one has a rich class of controlled Lagrangians to work with. Our theory provides such a class.*

This approach is attractive because it is done within the context of mechanics; one can understand the stabilization in terms of the effective creation of an energy extremum.

One can still ask many questions about this construction, such as: the role of damping, the swing-up problem, the efficiency and energy consumption of the method, etc. These issues are not all settled.

For problems with nonholonomic constraints (like a bicycle) there is reason to believe that a similar construction will work. This is based on recent advances in the geometry of nonholonomic systems and the associated stability theory.

A rigid body with a symmetric rotor. Now we give another stabilization problem that is solved by the same technique as the inverted pendulum. This example, investigated by Bloch, Krishnaprasad, Marsden and Sanchez [1992], provided a main motivation for the general approach of controlled Lagrangians.

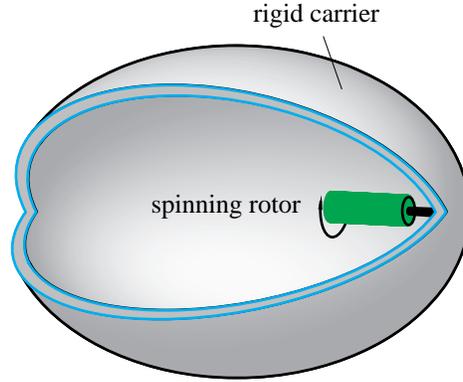


Figure 4.3: A rigid body with a rotor aligned on the long axis.

Consider a carrier rigid body with a rotor aligned along the third principal axis. The rotor spins under the influence of a torque u , as in figure 4.3.

The *equations of motion* are

$$\dot{\Pi} = \Pi \times \Omega, \quad \dot{l} = u$$

where $I_1 > I_2 > I_3$ are the carrier *moments of inertia*, $J_1 = J_2$ and J_3 are the rotor *moments of inertia*, $\Omega = (\Omega_1, \Omega_2, \Omega_3)$ are the carrier *angular velocity*, and α is the *relative angle of the rotor*. The *body angular momenta* are given by

$$\begin{aligned} \Pi_1 &= \lambda_1 \Omega_1; & \Pi_2 &= \lambda_2 \Omega_2 \\ \Pi_3 &= \lambda_3 \Omega_3 + J_3 \dot{\alpha}; & l_3 &= J_3(\Omega_3 + \dot{\alpha}) \end{aligned}$$

where $\lambda_i = I_i + J_i$.

The *equations in components* read

$$\begin{aligned} \dot{\Pi}_1 &= \left(\frac{1}{I_3} - \frac{1}{\lambda_2} \right) \Pi_2 \Pi_3 - \frac{l_3 \Pi_2}{I_3} \\ \dot{\Pi}_2 &= \left(\frac{1}{\lambda_1} - \frac{1}{I_3} \right) \Pi_1 \Pi_3 + \frac{l_3 \Pi_1}{I_3} \\ \dot{\Pi}_3 &= \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1} \right) \Pi_1 \Pi_2 \\ \dot{l}_3 &= u. \end{aligned}$$

If $u = 0$, then l_3 is a constant of motion and the remaining equations are Hamiltonian (Lie-Poisson) with

$$H = \frac{1}{2} \left(\frac{\Pi_1^2}{\lambda_1} + \frac{\Pi_2^2}{\lambda_2} + \frac{(\Pi_3 - l_3)^2}{I_3} \right) + \frac{1}{2} l_3^2.$$

We use the *feedback control law*:

$$u = k \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1} \right) \Pi_1 \Pi_2,$$

where k is a *gain parameter*. The system retains the S^1 symmetry and $P_k = l_3 - k\Pi_3$ is a new conserved quantity. The *closed loop equations* (eliminating the rotor variable) are

$$\begin{aligned}\dot{\Pi}_1 &= \Pi_2 \left(\frac{(1-k)\Pi_3 - P_k}{I_3} \right) - \frac{\Pi_3\Pi_2}{\lambda_2} \\ \dot{\Pi}_2 &= -\Pi_1 \left(\frac{(1-k)\Pi_3 - P_k}{I_3} \right) + \frac{\Pi_1\Pi_3}{\lambda_1} \\ \dot{\Pi}_3 &= \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1} \right) \Pi_1\Pi_2.\end{aligned}$$

These equations are still Hamiltonian with

$$H = \frac{1}{2} \left(\frac{\Pi_1^2}{\lambda_1} + \frac{\Pi_2^2}{\lambda_2} + \frac{((1-k)\Pi_3 - P_k)^2}{(1-k)I_3} \right) + \frac{1}{2} \frac{P_k^2}{J_3(1-k)},$$

using the Lie-Poisson (rigid body) Poisson structure on $so(3)^*$.

Noteworthy special cases are

1. $k = 0$, the *uncontrolled case*,
2. $k = J_3/\lambda_3$, the *driven case*, where $\dot{\alpha} = \text{constant}$.

How this fits into the general scheme of controlled Lagrangians. Start with the free Lagrangian:

$$L_0 = \frac{1}{2}(\lambda_1\Omega_1^2 + \lambda_2\Omega_2^2) + \frac{1}{2}J_3\Omega_3^2 + \frac{1}{2}J_3(\Omega_3 + \dot{\alpha})^2.$$

Consider the conserved quantity

$$P_0 = J_3(\Omega_3 + \dot{\alpha}) = l_3$$

associated with the S^1 action (rotor symmetry). Choose a horizontal one form to *change the connection*:

$$\tau = r\Omega_3$$

for a suitable constant r . Now construct a *new Lagrangian* obtained by replacing $\dot{\alpha}$ by $\dot{\alpha} + \tau_Q$ and modifying the metrics on the horizontal and vertical spaces using two scalars σ and ρ to produce a controlled Lagrangian $L_{\tau,\sigma,\rho}$.

With suitable r, σ, ρ , the momentum conjugate to α for this Lagrangian is P_k (up to a factor) and the resulting Euler-Poincaré equations give the feedback controlled system! Thus, our construction explains the otherwise “strange” Lagrangian and Hamiltonian structures found earlier by hand. This construction also works for the stabilization problem for *underwater vehicle dynamics*, discussed in Lecture 2. For details, see Bloch, Leonard and Marsden [1998].

Stabilization. With the control problem in *Hamiltonian form*, we can use the energy-Casimir method for stability. Let $P = 0$ and consider the equilibrium $(0, M, 0)$. For $k > 1 - J_3/\lambda_2$, this equilibrium is stable. This is proved by considering $H + C$ where $C = \varphi(\|\Pi\|^2)$. Pick φ so that

$$\delta(H + C)|_{(0,M,0)} = 0.$$

One computes that $\delta^2(H + C)$ is negative definite if $k > 1 - J_3/\lambda_2$ and $\varphi''(M^2) < 0$.

The phase portrait is that of the standard rigid body for the uncontrolled case where $k = 0$ (see Lecture 1). The feedback control in effect modifies the Lagrangian to *interchange the moments of inertia* of the system. The stabilization that takes place as the gain is increased can be viewed in terms of a *modification of the phase portrait of the rigid body*. See Figure 4.4.

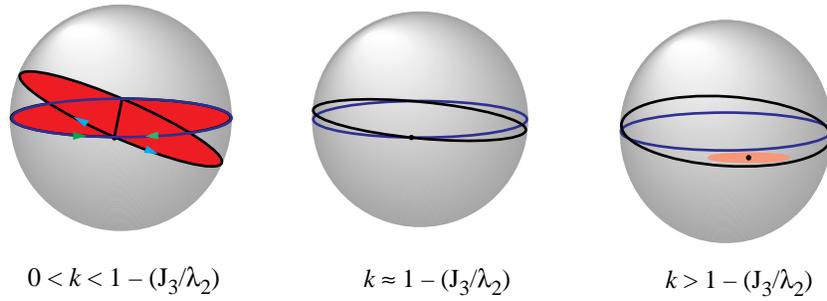


Figure 4.4: Stabilization by feedback.

Lecture 5

Variational Integrators

For conservative mechanical systems with symmetry, it is of interest to develop numerical schemes that preserve this symmetry, so that the associated conserved quantities are preserved *exactly* by the integration process. One would also like the algorithm to preserve either the Hamiltonian or the symplectic structure—one cannot expect to do *both* in general, as we shall show below. There is numerical evidence¹ that these mechanical integrators perform especially well for long time integrations, in which chaotic dynamics can be expected. Some standard algorithms can introduce spurious effects (such as nonexistent chaos or incorrect dissipation) in long integration runs.² We use the term *mechanical integrator* for an algorithm that respects one or more of the fundamental properties of being symplectic, preserving energy, or preserving the momentum map.

Symplectic integrators have been successfully used in a wide range of interesting applications from molecular dynamics to integrations of the solar system.

Basic definitions and examples. By an *algorithm* on a phase space P we mean a collection of maps $F_\tau : P \rightarrow P$ (usually depending smoothly on $\tau \in \mathbb{R}$ for small $\tau > 0$ and $z \in P$). Sometimes we write $z^{k+1} = F_\tau(z^k)$ for the algorithm and we write Δt or h for the *step size* τ . We say that the algorithm is *consistent* or is *first order accurate* with a vector field X on P if

$$\left. \frac{d}{d\tau} F_\tau(z) \right|_{\tau=0} = X(z). \quad (5.1)$$

Higher order accuracy is defined similarly by matching higher order derivatives. We say that the algorithm is *convergent* when

$$\lim_{n \rightarrow \infty} (F_{t/n})^n(z) = \varphi_t(z) \quad (5.2)$$

where φ_t is the flow of X . There are some general theorems guaranteeing convergence, with an important hypothesis being *stability*; *i.e.*, $(F_{t/n})^n(z)$ must remain close to z for small t and all $n = 1, 2, \dots$. We refer to Chorin, Hughes, Marsden and McCracken [1978] and Abraham, Marsden and Ratiu [1988] for details. An example of this is the (*Lie-Trotter*) *product formula*

$$e^{t(A+B)} = \lim_{n \rightarrow \infty} (e^{tA/n} e^{tB/n})^n \quad (5.3)$$

¹See, for example Channell and Scovel [1990], Marsden, O'Reilly, Wicklin and Zombro [1991], Pullin and Saffman [1991] and articles in Marsden, Patrick and Shadwick [1996].

²See, for example, Reinhall, Caughey, and Storti [1989].

which is a time splitting method for solving $\dot{x} = Ax + Bx$ by iterating, in an alternating fashion, the solutions of $\dot{x} = Ax$ and $\dot{x} = Bx$.

An algorithm F_τ is

1. a *symplectic integrator* if each F_τ is symplectic,
2. an *energy integrator* if $H \circ F_\tau = H$ (where $X = X_H$),
3. a *momentum integrator* if $\mathbf{J} \circ F_\tau = \mathbf{J}$, where \mathbf{J} is the momentum map for the action of a Lie group G .

If F_τ has one or more of these properties, we call it a *mechanical integrator*. Notice that if an integrator has one of these three properties, then so does any iterate of it, since these properties are preserved by composition.

There are several ways that have been employed to find mechanical integrators. For example, one can search amongst existing algorithms and find ones with special algebraic properties that make them symplectic or energy-preserving. Alternatively, one can attempt to design mechanical integrators from scratch. Here are some simple examples:

Example 1. A first order explicit symplectic scheme in the plane is given by the map $(q_0, p_0) \mapsto (q, p)$ defined by

$$\begin{aligned} q &= q_0 + (\Delta t)p_0 \\ p &= p_0 - (\Delta t)V'(q_0 + (\Delta t)p_0). \end{aligned} \tag{5.4}$$

This map is a first order approximation to the flow of Hamilton's equations for the Hamiltonian $H = (p^2/2) + V(q)$. Here, one can verify by direct calculation that this scheme is symplectic. \blacklozenge

Example 2. An implicit, symplectic, second order accurate scheme in the plane for the same Hamiltonian as in Example 1 is

$$\begin{aligned} q &= q_0 + (\Delta t)(p + p_0)/2 \\ p &= p_0 - (\Delta t)V'((q + q_0)/2). \end{aligned} \quad \blacklozenge \tag{5.5}$$

Shortly we shall see how to construct such algorithms systematically, but for now just regard them as illustrative. The next example shows that the second order accurate mid-point rule is symplectic (Feng [1986]). This algorithm is also useful in developing almost Poisson integrators (Austin, Krishnaprasad and Wang [1993]).

Example 3. In a symplectic vector space the following *mid point rule* for a Hamiltonian vector field X_H is symplectic:

$$\frac{z^{k+1} - z^k}{\Delta t} = X_H \left(\frac{z^k + z^{k+1}}{2} \right). \tag{5.6}$$

Notice that for small Δt the map given implicitly by this equation is well defined by the implicit function theorem. One way to show that it is symplectic, is to use the fact that the *Cayley transform* S of an infinitesimally symplectic linear map A , namely

$$S = (\mathbf{1} - \lambda A)^{-1} (\mathbf{1} + \lambda A) \tag{5.7}$$

is symplectic if $\mathbf{1} - \lambda A$ is invertible for some real λ . To apply this to our situation, rewrite the algorithm (5.6) as

$$F_\tau(z) - z - \tau X_H \left(\frac{z + F_\tau(z)}{2} \right) = 0. \quad (5.8)$$

Letting $S = \mathbf{D}F_\tau(z)$ and $A = \mathbf{D}X_H \left(\frac{z + F_\tau(z)}{2} \right)$ we get, by differentiation of equation (5.8) with respect to z , $S - \mathbf{1} - \frac{1}{2}\tau A(\mathbf{1} + S) = 0$; *i.e.*, (5.7) holds with $\lambda = \tau/2$. Thus, (5.6) defines a symplectic scheme. \blacklozenge

Example 4. Here is an example of an implicit energy preserving algorithm from Chorin, Hughes, Marsden and McCracken [1978]. Consider a Hamiltonian system for $\mathbf{q} \in \mathbb{R}^n$ and $\mathbf{p} \in \mathbb{R}^n$:

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}. \quad (5.9)$$

Define the following implicit scheme

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \frac{H(\mathbf{q}_{n+1}, \mathbf{p}_{n+1}) - H(\mathbf{q}_{n+1}, \mathbf{p}_n)}{\lambda^T(\mathbf{p}_{n+1} - \mathbf{p}_n)} \lambda, \quad (5.10)$$

$$\mathbf{p}_{n+1} = \mathbf{p}_n - \Delta t \frac{H(\mathbf{q}_{n+1}, \mathbf{p}_n) - H(\mathbf{q}_n, \mathbf{p}_n)}{\mu^T(\mathbf{q}_{n+1} - \mathbf{q}_n)} \mu, \quad (5.11)$$

where

$$\lambda = \frac{\partial H}{\partial \mathbf{p}}(\alpha \mathbf{q}_{n+1} + (1 - \alpha) \mathbf{q}_n, \beta \mathbf{p}_{n+1} + (1 - \beta) \mathbf{p}_n), \quad (5.12)$$

$$\mu = \frac{\partial H}{\partial \mathbf{q}}(\gamma \mathbf{q}_{n+1} + (1 - \gamma) \mathbf{q}_n, \delta \mathbf{p}_{n+1} + (1 - \delta) \mathbf{p}_n), \quad (5.13)$$

and where $\alpha, \beta, \gamma, \delta$ are arbitrarily chosen constants in $[0, 1]$.

The proof of conservation of energy is simple: From (5.10), we have

$$(\mathbf{q}_{n+1} - \mathbf{q}_n)^T (\mathbf{p}_{n+1} - \mathbf{p}_n) = \Delta t (H(\mathbf{q}_{n+1}, \mathbf{p}_{n+1}) - H(\mathbf{q}_{n+1}, \mathbf{p}_n)), \quad (5.14)$$

and from (5.11)

$$(\mathbf{p}_{n+1} - \mathbf{p}_n)^T (\mathbf{q}_{n+1} - \mathbf{q}_n) = -\Delta t (H(\mathbf{q}_{n+1}, \mathbf{p}_n) - H(\mathbf{q}_n, \mathbf{p}_n)). \quad (5.15)$$

Subtracting (5.15) from (5.14), we obtain

$$H(\mathbf{q}_{n+1}, \mathbf{p}_{n+1}) = H(\mathbf{q}_n, \mathbf{p}_n). \quad (5.16)$$

This algorithm is checked to be consistent. In general, it is not symplectic. \blacklozenge

Example 5. Let us apply the product formula idea to the simple pendulum. The equations are

$$\frac{d}{dt} \begin{pmatrix} \varphi \\ p \end{pmatrix} = \begin{pmatrix} p \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -\sin \varphi \end{pmatrix}.$$

Each vector field can be integrated explicitly to give maps

$$G_\tau(\varphi, p) = (\varphi + \tau p, p)$$

and

$$H_\tau(\varphi, p) = (\varphi, p - \tau \sin \varphi)$$

each of which is symplectic. Thus, the composition $F_\tau = G_\tau \circ H_\tau$, namely,

$$F_\tau(q, p) = (\varphi + \tau p - \tau^2 \sin \varphi, p - \tau \sin \varphi)$$

is a first order symplectic scheme for the simple pendulum. It is closely related to the **standard map**. The orbits of F_τ need not preserve energy and they may be chaotic, while the trajectories of the simple pendulum are of course not chaotic. ♦

We refer to the cited references, and to Ruth [1983], Feng [1986], Sanz-Serna [1988], Sanz-Serna and Calvo [1994] and references therein for more examples, including symplectic Runge-Kutta schemes. Product formula techniques are also discussed in McLachlan and Scovel [1996].

Limitations on mechanical integrators. A number of algorithms have been developed specifically for integrating Hamiltonian systems to conserve the energy integral, but without attempting to capture all of the details of the Hamiltonian structure (see Example 4 above and also Stofer [1987] and Greenspan [1974, 1984]). In fact, some of the standard energy-conservative algorithms have poor momentum behavior over moderate time ranges. This makes them unsuitable for problems in satellite dynamics for example, where the exact conservation of a momentum integral is essential to problems in attitude control.

One can get angular momentum drift in energy-conservative simulations of, for example, rods that are free to vibrate and rotate. To control such drifts and attain the high levels of computational accuracy demanded by automated control mechanisms, one would be forced to reduce computational step sizes to such an extent that the numerical simulation would be prohibitively inefficient. Similarly, if one attempts to use a standard energy-conservative algorithm to simulate both the rotational and vibrational modes of a freely moving flexible rod, the algorithm may predict that the rotational motion will come to a virtual halt after only a few cycles! For a documented simulation of a problem with momentum conservation, see Simo and Wong [1989]. Unless one designs the algorithms carefully, in the process of enforcing energy conservation, one could upset conservation of angular momentum.

What may seem surprising is that all of the implicit members of the Newmark family, perhaps the most widely used time-stepping algorithms in nonlinear structural dynamics, *are not designed to conserve energy and also fail to conserve momentum*. Among the explicit members, only the central difference method preserves momentum. The proof of these results is in Simo, Tarnow and Wong [1991]; in fact, central difference schemes are variational in the sense of Veselov which we shall explain below, and this may be viewed as one explanation for why they preserve momentum.

In traditional integrators, much attention has been paid to energy conservation properties, some, as we have noted to momentum conservation, and even less to conserving the symplectic or Poisson structure. However, one can imagine that it is also quite important.

Given the importance of conserving integrals of motion and the important role played by the Hamiltonian structure in the reduction procedure for a system with symmetry, one might hope to find an algorithm that combines all of the desirable properties: conservation of energy, conservation of momenta (and other independent integrals), and conservation of the symplectic structure. However, one cannot do all three of these things at once unless one relaxes one or more of the conditions in the following sense:

Proposition 5.1 (Ge and Marsden [1988]) *If an algorithm for a given Hamiltonian system X_H with a symmetry group G is energy preserving, symplectic, momentum preserving and G -equivariant and if the dynamics of X_H is nonintegrable on the reduced space (in the sense spelled out in the proof) then the algorithm already gives the exact solution of the given problem up to a time reparametrization.*

Proof Suppose $F_{\Delta t}$ is our symplectic algorithm of the type discussed above, and consider the application of the algorithm to the reduced phase space. We assume that the Hamiltonian H is the only integral of motion of the reduced dynamics (*i.e.*, all other integrals of the system have been found and taken out in the reduction process in the sense that any other conserved quantity (in a suitable class) is functionally dependent on H). Since $F_{\Delta t}$ is symplectic it is the Δt -time map of some *time-dependent* Hamiltonian function K . Now assume that the symplectic map $F_{\Delta t}$ also conserves H for *all* values of Δt . Thus $\{H, K\} = 0 = \{K, H\}$. The latter equation implies that K is functionally dependent on H since the flow of H (the “true dynamics”) had no other integrals of motion. The functional dependence of K on H in turn implies that their Hamiltonian vector fields are parallel, so the flow of K (the approximate solution) and the flow of H (the exact solution) must lie along identical curves in the reduced phase space; thus the flows are equivalent up to time reparametrization. ■

Colloqually speaking, this result means that it is unlikely one can find an algorithm that simultaneously conserves the symplectic structure, the momentum map, and the Hamiltonian. It is tempting (but probably wrong) to guess from this that one can monitor accuracy by keeping track of all three. However, for small time steps, symplectic integrators have *surrogate Hamiltonians* as shown by Neistadt [1984]; one gets nearby exactly conserved energy functions with errors that are exponentially small (presumably below round off errors) in the time step.

Non-symplectic algorithms that conserve both momentum and energy have been studied by Simo and Tarnow [1992], Simo and Wong [1989] and Austin, Krishnaprasad and Wang [1993]. Dissipative effects can often be dealt with by means of *product formulas*. See Armero and Simo [1992, 1993, 1996] for example.

Variational methods. Symplectic-momentum integrators can be simply and naturally constructed by means of *discretizations of Hamilton’s principle* following ideas of Veselov [1988, 1991]. We shall explain this procedure following the exposition of Wendlandt and Marsden [1997].

The emphasis here is on *simplicity*: the theory and practice are easy! For example, they can be much simpler to use and implement than generating function methods (although the two techniques are closely related theoretically).

Variational integrators *include the popular Verlet methods and shake algorithms* as special cases. The methods also *handle constraints* in a simple way—this is one of the sterling features of the variational technique.

The methods generalize to pde’s using multisymplectic geometry with the result being a class of *multisymplectic momentum integrators*. One obtains, in a natural way,

spacetime integrators. See Marsden, Patrick and Shkoller [1998] for details and numerical examples.

There is still *lots to do* in this area! A goal is to continue developing the theory and to implement multisymplectic integrators for various interesting pde systems, such as those for nonlinear optics (the NLS equation), MHD, ocean dynamics, etc.

The discrete variational principle. We now discuss *discrete variational principles (DVP)* which lead to evolution algorithms analogous to the Euler-Lagrange equations from mechanics. As we shall see, an algorithm is implicitly determined by the associated *discrete Euler-Lagrange (DEL)* equations. Let us now explain how this works.

Given a *configuration space* Q , a **discrete Lagrangian** is a map

$$\mathbb{L} : Q \times Q \rightarrow \mathbb{R}.$$

In practice, \mathbb{L} is obtained by approximating a given Lagrangian, but regard \mathbb{L} as given for the moment. The time step information will be contained in \mathbb{L} .

For a positive integer N , the **action sum** is the map $\mathbb{S} : Q^{N+1} \rightarrow \mathbb{R}$ defined by

$$\mathbb{S} = \sum_{k=0}^{N-1} \mathbb{L}(q_{k+1}, q_k),$$

where $q_k \in Q$ and k is a nonnegative integer. The action sum is the discrete analog of the action integral in mechanics.

The **discrete variational principle** states that the evolution equations extremize the action sum given fixed end points, q_0 and q_N . Extremizing \mathbb{S} over q_1, \dots, q_{N-1} leads to the **DEL equations**:

$$D_2\mathbb{L}(q_{k+1}, q_k) + D_1\mathbb{L}(q_k, q_{k-1}) = 0$$

for all $k = 1, \dots, N-1$. We can write this equation in terms of a discrete **algorithm**

$$\Phi : Q \times Q \rightarrow Q \times Q$$

defined implicitly by

$$D_2\mathbb{L} \circ \Phi + D_1\mathbb{L} = 0,$$

i.e.,

$$\Phi(q_k, q_{k-1}) = (q_{k+1}, q_k).$$

If, for each $q \in Q$, $D_2\mathbb{L}(q, q) : T_q Q \rightarrow T_q^* Q$ is invertible, then $D_2\mathbb{L} : Q \times Q \rightarrow T^* Q$ is locally invertible and so the algorithm Φ , which flows the system forward in discrete time, is well defined for small time steps.

In coordinates, q^i on Q , the DEL equations are

$$\frac{\partial \mathbb{L}}{\partial q_k^i} \circ \Phi(q_{k+1}, q_k) + \frac{\partial \mathbb{L}}{\partial q_{k+1}^i}(q_{k+1}, q_k) = 0$$

i.e.,

$$\frac{\partial \mathbb{L}}{\partial q_{k+1}^i}(q_{k+2}, q_{k+1}) + \frac{\partial \mathbb{L}}{\partial q_{k+1}^i}(q_{k+1}, q_k) = 0.$$

The algorithm is symplectic. To explain the sense in which the algorithm is symplectic, first define the *fiber derivative* by

$$\mathbb{F}\mathbb{L} : Q \times Q \rightarrow T^*Q; (q_1, q_0) \mapsto (q_0, D_2\mathbb{L}(q_1, q_0))$$

and define the 2-form ω on $Q \times Q$ by pulling back the canonical 2-form on T^*Q :

$$\omega = \mathbb{F}\mathbb{L}^*(\Omega_{\text{CAN}}).$$

The fiber derivative is analogous to the standard Legendre transform. The coordinate expression for ω is:

$$\omega = \frac{\partial^2 \mathbb{L}}{\partial q_k^i \partial q_{k+1}^j} (q_{k+1}, q_k) dq_k^i \wedge dq_{k+1}^j.$$

Theorem 5.2 *The algorithm Φ exactly preserves the symplectic form ω .*

One proof of this is to simply verify it with a straightforward calculation—see Wendlandt and Marsden [1997] for the details. Another is to derive this information directly from the variational structure. We will come back to this shortly.

The algorithm preserves momentum. Recall that *Noether's theorem* states that a continuous symmetry of the Lagrangian leads to conserved quantities, as with linear and angular momentum. A nice way to derive these conservation laws (the way Noether did it) is to use the *invariance of the variational principle*.

Let the discrete Lagrangian be invariant under the action of a Lie group G on Q , and let $\xi \in \mathfrak{g}$, the Lie algebra of G . By analogy with the continuous case, define the *discrete momentum map*, $\mathbb{J} : Q \times Q \rightarrow \mathfrak{g}^*$ by

$$\langle \mathbb{J}(q_{k+1}, q_k), \xi \rangle := \langle D_2\mathbb{L}(q_{k+1}, q_k), \xi_Q(q_k) \rangle,$$

Theorem 5.3 *The algorithm Φ exactly preserves the momentum map.*

The discrete momentum map \mathbb{J} is equivariant with respect to the action of G on $Q \times Q$ and the coadjoint action of G on \mathfrak{g}^* . We also note that the algorithm Φ can be obtained by using $-\mathbb{L}$ as a generating function. Again, one can check this by a calculation or by invoking invariance of the variational principle—which we will indicate below.

Construction of mechanical integrators. We show how to construct integrators in a practical manner by *enforcing the constraints through Lagrange multipliers*. (Wendlandt and Marsden [1997] also discuss the Jacobian used to solve the nonlinear equations; the constrained coordinate formulation has a special structure that can be exploited to increase simulation efficiency, as well as local truncation error and solvability.)

Assume that we have a mechanical system with a constraint manifold, $Q \subset V$, where V is a real finite dimensional vector space, and that we have an unconstrained Lagrangian, $L : TV \rightarrow \mathbb{R}$ which, by restriction of L to TQ , defines a *constrained Lagrangian*, $L^c : TQ \rightarrow \mathbb{R}$. Roughly speaking, V is a containing vector space in which the computer arithmetic will take place. In particular, coordinate charts on Q are *not* chosen for this purpose. In fact, apart from the use of the containing vector space V , the algorithms developed here are independent of the use of coordinates on Q .

We also assume that we have a vector valued *constraint function*, $g : V \rightarrow \mathbb{R}^k$, such that our constraint manifold is given by $g^{-1}(0) = Q \subset V$, with 0 a regular value of g . The dimension of V is denoted n , and therefore, the dimension of Q is $m = n - k$.

Define a **discrete, unconstrained Lagrangian**, $\mathbb{L} : V \times V \rightarrow \mathbb{R}$ by

$$\mathbb{L}(y, x) = L\left(\frac{y+x}{2}, \frac{y-x}{h}\right),$$

where $h \in \mathbb{R}_+$ is the time step. (This is not the only possible choice but it is one that leads to a second order accurate algorithm.) The **unconstrained action sum** is defined by

$$\mathbb{S} = \sum_{k=0}^{N-1} \mathbb{L}(v_{k+1}, v_k).$$

Extremize $\mathbb{S} : V^{N+1} \rightarrow \mathbb{R}$ subject to the constraint that $v_k \in Q \subset V$ for $k = 1, \dots, N-1$, i.e., solve

$$D_2\mathbb{L}(v_{k+1}, v_k) + D_1\mathbb{L}(v_k, v_{k-1}) + \lambda_k^T Dg(v_k) = 0$$

(no sum on k) with $g(v_k) = 0$ for $k = 1, \dots, N-1$. Here λ_k are *Lagrange multipliers*, chosen to enforce the constraints.

Summary. The **algorithm** is defined by starting with v_k and v_{k-1} in $Q \subset V$, i.e., $g(v_k) = 0$ and $g(v_{k-1}) = 0$, and solving

$$D_2\mathbb{L}(v_{k+1}, v_k) + D_1\mathbb{L}(v_k, v_{k-1}) + \lambda_k^T Dg(v_k) = 0$$

subject to $g(v_{k+1}) = 0$, for v_{k+1} and λ_k .

In terms of the *unconstrained Lagrangian*, the algorithm reads as follows:

$$\begin{aligned} & \frac{1}{h} \left[\frac{\partial L}{\partial \dot{v}} \left(\frac{v_k + v_{k-1}}{2}, \frac{v_k - v_{k-1}}{h} \right) - \frac{\partial L}{\partial \dot{v}} \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right) \right] \\ & + \frac{1}{2} \left[\frac{\partial L}{\partial v} \left(\frac{v_k + v_{k-1}}{2}, \frac{v_k - v_{k-1}}{h} \right) + \frac{\partial L}{\partial v} \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right) \right] \\ & + D^T g(v_k) \lambda_k = 0 \end{aligned}$$

together with $g(v_{k+1}) = 0$.

Example. If the continuous Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^T M \dot{q} - V(q)$$

with constraint $g(q) = 0$, where M is a constant mass matrix, and V is the potential energy, then the DEL equations are

$$\begin{aligned} & M \left(\frac{v_{k+1} - 2v_k + v_{k-1}}{h^2} \right) + \frac{1}{2} \left(\frac{\partial V}{\partial q} \left(\frac{v_{k+1} + v_k}{2} \right) + \frac{\partial V}{\partial q} \left(\frac{v_k + v_{k-1}}{2} \right) \right) \\ & - D^T g(v_k) \lambda_k = 0 \end{aligned}$$

with $g(v_{k+1}) = 0$.

These algorithms produce the Verlet and shake algorithms as special cases, as is discussed in Wendlandt and Marsden [1997]. The following result is very plausible: it states that the earlier intrinsic construction is coincident with this construction using constraints.

Theorem 5.4 *The algorithm defined using Lagrange multipliers coincides with that defined intrinsically using the constrained discrete Lagrangian on $Q \times Q$, so it is symplectic and momentum preserving.*

The examples treated in Wendlandt and Marsden [1997] are simple ones, the double spherical pendulum (dsp) and the rigid body. The former example is interesting because of its *chaotic dynamics and pattern evocation phenomenon* (Marsden and Scheurle [1995]).

Both are constrained systems and may be readily integrated by the above methods. Numerical issues such as cpu time and accuracy are also discussed, as well as the comparison of these methods with energy-momentum methods. For the rigid body, we use quaternions to handle the constraints but this was just for illustration and of course one could handle the constraints directly using $SO(3) \subset GL(3)$ as well. The following figure from one of the dsp calculations shows the surrogate energy phenomenon.

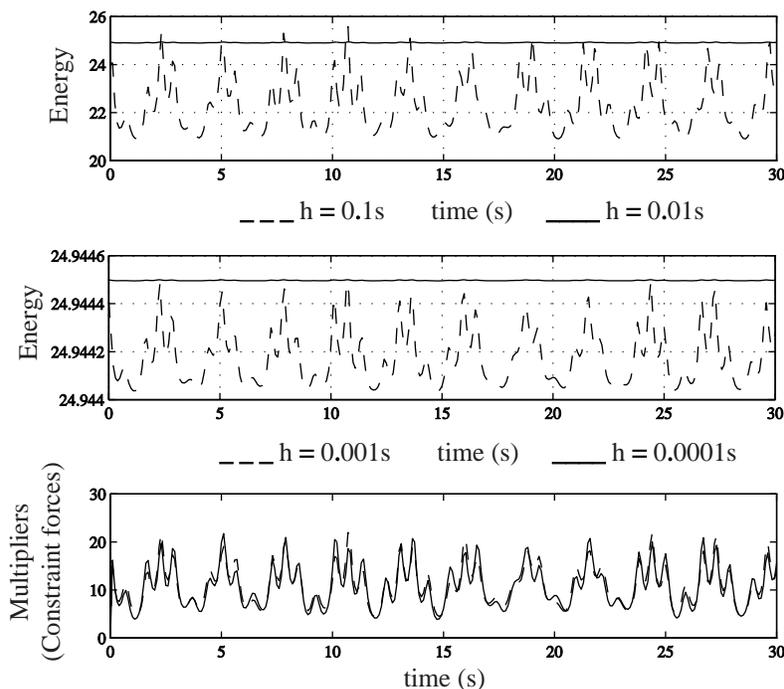


Figure 5.1: Energy and multipliers versus time for the dsp simulation.

An intrinsic variational viewpoint. Recall that given a Lagrangian function $L : TQ \rightarrow \mathbb{R}$, we construct the corresponding *action functional* \mathfrak{S} on C^2 curves $q(t)$ by (using coordinate notation)

$$\mathfrak{S}(q(\cdot)) \equiv \int_a^b L \left(q^i(t), \frac{dq^i}{dt}(t) \right) dt. \quad (5.17)$$

The action functional depends on a and b , but this is not explicit in the notation. Hamilton's principle seeks the curves $q(t)$ for which the functional \mathfrak{S} is stationary under variations of $q^i(t)$ with fixed endpoints. It will be useful to recall this calculation; namely, we seek curves

$q(t)$ which satisfy

$$\mathbf{d}\mathfrak{S}(q(t)) \cdot \delta q(t) \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \mathfrak{S}(q(t) + \epsilon \delta q(t)) = 0 \quad (5.18)$$

for all $\delta q(t)$ with $\delta q(a) = \delta q(b) = 0$. Abbreviating $q_\epsilon \equiv q + \epsilon \delta q$, and using integration by parts, the calculation is

$$\begin{aligned} \mathbf{d}\mathfrak{S}(q(t)) \cdot \delta q(t) &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \int_a^b L \left(q_\epsilon^i(t), \frac{dq_\epsilon^i}{dt}(t) \right) dt \\ &= \int_a^b \delta q^i \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) dt + \left. \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right|_a^b. \end{aligned} \quad (5.19)$$

The last term in (5.19) vanishes since $\delta q(a) = \delta q(b) = 0$, so that the requirement (5.18) for \mathfrak{S} to be stationary yields the **Euler-Lagrange equations**

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = 0. \quad (5.20)$$

Recall that L is called **regular** when the matrix $[\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j]$ is everywhere a nonsingular matrix and in this case, the Euler-Lagrange equations are second order ordinary differential equations.

Since the action (5.17) is independent of the choice of coordinates, the Euler-Lagrange equations are coordinate independent as well. Consequently, it is natural that the Euler-Lagrange equations may be intrinsically expressed using the language of differential geometry. This intrinsic development of mechanics is now standard, and can be seen, for example, in Arnold [1989], Abraham and Marsden [1978], and Marsden and Ratiu [1998].

The **canonical 1-form** on the $2n$ -dimensional cotangent bundle T^*Q is defined by

$$\Theta_0(\alpha_q)w_{\alpha_q} \equiv \alpha_q \cdot T\pi_Q w_{\alpha_q},$$

where $\alpha_q \in T_q^*Q$, and $w_{\alpha_q} \in T_{\alpha_q}T^*Q$, and $\pi_Q : T^*Q \rightarrow Q$ is the projection. The Lagrangian L intrinsically defines a fiber preserving bundle map $\mathbb{F}L : TQ \rightarrow T^*Q$, the **Legendre transformation**, by vertical differentiation:

$$\mathbb{F}L(v_q)w_q \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} L(v_q + \epsilon w_q).$$

One normally defines the **Lagrange 1-form** on TQ by pull-back $\Theta_L \equiv \mathbb{F}L^*\Theta_0$, and the **Lagrange 2-form** by $\Omega_L = -\mathbf{d}\Theta_L$. We then seek a vector field X_E (called the **Lagrange vector field**) on TQ such that $X_E \lrcorner \Omega_L = dE$, where the **energy** E is defined by $E(v_q) \equiv \mathbb{F}L(v_q)v_q - L(v_q)$. A number of these definitions can be given a little more directly on TQ , without resorting to T^*Q , but it amounts to the same thing.

If $\mathbb{F}L$ is a local diffeomorphism then X_E exists and is unique, and its integral curves solve the Euler-Lagrange equations. In addition, the flow F_t of X_E is symplectic; that is, preserves Ω_L : $F_t^*\Omega_L = \Omega_L$. These facts are usually proved using differential forms and Lie derivatives.

Despite the compactness and precision of this differential-geometric approach, it is difficult to motivate and also is not entirely contained on the Lagrangian side. The canonical 1-form Θ_0 seems to appear from nowhere, as does the Legendre transform $\mathbb{F}L$.

The variational approach. Besides being more faithful to history, more and more, one is finding that there are advantages to staying on the “Lagrangian side”. Many examples can be given, but the theory of Lagrangian reduction discussed in lecture 1 (the Euler-Poincaré equations being an instance) and associated control issues discussed in lectures 3 and 4 are examples. Other examples are the direct variational approach to questions in black hole dynamics given by Wald [1993] and the development of variational asymptotics (see Holm [1996], Holm, Marsden and Ratiu [1998b], and references therein). In such studies, it is the variational principle that is the center of attention.

One can derive in a natural way the entire differential geometric structures, including momentum mappings, directly from the variational approach. This development begins by removing the endpoint condition $\delta q(a) = \delta q(b) = 0$. Equation (5.19) becomes

$$d\mathfrak{S}(q(\cdot)) \cdot \delta q(\cdot) = \int_a^b \delta q^i \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) dt + \left. \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right|_a^b, \quad (5.21)$$

but now the left side operates on more general δq and the last term on the right side need not vanish. That last term of (5.21) is a linear pairing of the function $\partial L / \partial \dot{q}^i$, a function of q^i and \dot{q}^i , with the tangent vector $\delta \dot{q}^i$. Thus, one may consider it a 1-form on TQ ; namely the 1-form $(\partial L / \partial \dot{q}^i) d\dot{q}^i$. This is exactly the Lagrange 1-form, and we can summarize this as follows:

Theorem 5.5 *Given a C^k Lagrangian L , $k \geq 2$, there exists a unique C^{k-2} mapping $D_{EL}L : \ddot{Q} \rightarrow T^*Q$, defined on the second order submanifold*

$$\ddot{Q} \equiv \left\{ \left. \frac{d^2 q}{dt^2}(0) \in T^2Q \right| q \text{ is a } C^2 \text{ curve in } Q \right\}$$

of T^2Q , and a unique C^{k-1} 1-form Θ_L on TQ , such that, for all C^2 variations $q_\epsilon(t)$,

$$d\mathfrak{S}(q(\cdot)) \cdot \delta q(\cdot) = \int_a^b D_{EL}L \left(\frac{d^2 q}{dt^2} \right) \cdot \delta q dt + \Theta_L \left(\frac{dq}{dt} \right) \cdot \hat{\delta} q \Big|_a^b, \quad (5.22)$$

where

$$\delta q(t) \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} q_\epsilon(t), \quad \hat{\delta} q(t) \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \left. \frac{d}{dt} \right|_{t=0} q_\epsilon(t).$$

The 1-form so defined is called the **Lagrange 1-form**.

Indeed, uniqueness and local existence are consequences of the calculation (5.19), and the coordinate independence of the action, and then global existence follows.

Thus, using the variational principle, the Lagrange 1-form Θ_L is the “boundary part” of the functional derivative of the action when the boundary is varied. The analogue of the symplectic form is the (negative of) exterior derivative of Θ_L ; i.e., $\Omega_L \equiv -d\Theta_L$.

Lagrangian flows are symplectic. One of Lagrange’s basic discoveries was that the solutions of the Euler-Lagrange equations give rise to a symplectic map. It is a curious twist of history that he did this without the machinery of either differential forms, of the Hamiltonian formalism or of Hamilton’s principle itself. (See Marsden and Ratiu [1998] for an account of this history.)

Assuming that L is regular, the variational principle gives as we have seen, coordinate independent second order ordinary differential equations. We temporarily denote the vector field on TQ so obtained by X , and its flow by F_t . Consider the restriction of \mathfrak{S} to the

subspace \mathcal{C}_L of solutions of the variational principle. The space \mathcal{C}_L may be identified with the initial conditions for the flow; to $v_q \in TQ$, we associate the integral curve $s \mapsto F_s(v_q)$, $s \in [0, t]$. The value of \mathfrak{S} on that curve is denoted by \mathfrak{S}_t , and again called the **action**. We thus regard \mathfrak{S}_t as a real valued function on TQ . The fundamental equation (5.22) becomes

$$\mathbf{d}\mathfrak{S}_t(v_q)w_{v_q} = \Theta_L(F_t(v_q)) \cdot \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} F_t(v_q + \epsilon w_{v_q}) - \Theta_L(v_q) \cdot w_{v_q},$$

where $\epsilon \mapsto v_q + \epsilon w_{v_q}$ symbolically represents any curve at v_q with derivative w_{v_q} . We have thus derived the equation

$$\mathbf{d}\mathfrak{S}_t = F_t^* \Theta_L - \Theta_L. \quad (5.23)$$

Taking the exterior derivative of (5.23) yields the fundamental fact that the flow of X is symplectic:

$$0 = \mathbf{d}\mathbf{d}\mathfrak{S}_t = \mathbf{d}(F_t^* \Theta_L - \Theta_L) = -F_t^* \Omega_L + \Omega_L$$

which is equivalent to $F_t^* \Omega_L = \Omega_L$. Thus, using the variational principle, the analogue that the evolution is symplectic is the equation $\mathbf{d}^2 = 0$, applied to the action restricted to space of solutions of the variational principle. Equation (5.23) also provides the differential-geometric equations for X . Indeed, one time derivative of (5.23) gives $\mathbf{d}L = \mathcal{L}_X \Theta_L$ so that

$$X \lrcorner \Omega_L = -X \lrcorner \mathbf{d}\Theta_L = -\mathcal{L}_X \Theta_L + \mathbf{d}(X \lrcorner \Theta_L) = \mathbf{d}(X \lrcorner \Theta_L - L) = \mathbf{d}E,$$

if we define $E \equiv X \lrcorner \Theta_L - L$. Thus, we quite naturally find that $X = X_E$.

Of course, this set up also leads directly to Hamilton-Jacobi theory, which was one of the ways in which symplectic integrators were developed (see McLachlan and Scovel [1996] and references therein.) However, we shall not pursue this aspect of the theory.

Momentum maps. Suppose that a Lie group G , with Lie algebra \mathfrak{g} , acts on Q , and hence on curves in Q , in such a way that the action \mathfrak{S} as defined by (5.17) is invariant. This is implied by, (but does not imply) that L itself is invariant³. Clearly, G sends solutions of the variational principle to themselves, so the action of G restricts to \mathcal{C}_L , and the action commutes with F_t . Denoting the infinitesimal generator of $\xi \in \mathfrak{g}$ on TQ by ξ_{TQ} , we have by (5.23),

$$0 = \xi_{TQ} \lrcorner \mathbf{d}\mathfrak{S}_t = \xi_{TQ} \lrcorner (F_t^* \Theta_L - \Theta_L) = F_t^* (\xi_{TQ} \lrcorner \Theta_L) - \xi_{TQ} \lrcorner \Theta_L. \quad (5.24)$$

For $\xi \in \mathfrak{g}$, define $J_\xi : TQ \rightarrow \mathbb{R}$ by $J_\xi \equiv \xi_{TQ} \lrcorner \Theta_L$. Then (5.24) says the J_ξ is an integral of the flow of X_E . We have arrived at a version of Noether's theorem (actually rather close to the original derivation of Noether): *Using the variational principle, Noether's theorem results from the infinitesimal invariance of the action restricted to space of solutions of the variational principle. The conserved momentum associated to a Lie algebra element ξ is $J_\xi = \xi \lrcorner \Theta_L$, where Θ_L is the Lagrange one-form.*

Reformulation in terms of first variations. We have seen that symplecticity of the flow and Noether's theorem result from restricting the action to the space of solutions. The tacit assumption is that the space of solutions is a manifold in some appropriate sense. This is a potential problem, since solution spaces for field theories are known to have singularities

³The distinction is sometimes important. See Olver [1986] for a discussion.

(see, eg, Arms, Marsden and Moncrief [1982]). We now show how this problem can be avoided.

Given a solution $q(t) \in \mathcal{C}_L$, a **first variation** at $q(t)$ is a vector field V on Q along $q(t)$ that is the derivative of some curve in \mathcal{C}_L at $q(t)$. When \mathcal{C}_L is a manifold, a first variation is just a vector at $q(t)$ tangent to \mathcal{C}_L . Temporarily define $\alpha \equiv \mathbf{d}\mathfrak{S} - \Theta_L$ where by abuse of notation Θ_L is the one form on \mathcal{C} defined by

$$\Theta_L(q(t))\delta q(t) \equiv \Theta_L(b)\delta q(b) - \Theta_L(a)\delta q(a).$$

Then \mathcal{C}_L is defined by $\alpha = 0$ and we have the equation

$$\mathbf{d}\mathfrak{S} = \alpha + \Theta_L,$$

so if V and W are first variations at $q(t)$, we obtain

$$0 = V \lrcorner W \lrcorner \mathbf{d}^2\mathfrak{S} = V \lrcorner W \lrcorner \mathbf{d}\alpha + V \lrcorner W \lrcorner \mathbf{d}\Theta_L. \quad (5.25)$$

We have the identity

$$\mathbf{d}\alpha(V, W)(q(t)) = V(\alpha(W)) - W(\alpha(V)) - \alpha([V, W]), \quad (5.26)$$

which we will use to evaluate (5.25) at the curve $q(t)$. Let $q_\epsilon^V(t)$ be a family of solutions tangent to V at $\epsilon = 0$. For the first term of (5.26), we have

$$V(\alpha(W))(q(t)) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \alpha(W)(q_\epsilon^V),$$

which vanishes, since α is zero along q_ϵ^V for every ϵ . Similarly the second term of (5.25) at $q(t)$ also vanishes, while the third term of vanishes since $\alpha(q(t)) = 0$. Consequently, symplecticity of the the Lagrangian flow F_t may be written

$$V \lrcorner W \lrcorner \mathbf{d}\Theta_L = 0,$$

for all first variations V and W . This formulation is valid whether or not the solution space is a manifold, and it does not explicitly refer to any temporal notion. Similarly, Noether's theorem may be written in this way. Summarizing, using the variational principle, the analogue that the evolution is symplectic is the equation $\mathbf{d}^2\mathfrak{S} = 0$ restricted to first variations of the space of solutions of the variational principle. The analogue of Noether's theorem is infinitesimal invariance of $\mathbf{d}\mathfrak{S}$ restricted to first variations of the space of solutions of the variational principle.

The variational route to the differential-geometric formalism has obvious pedagogical advantages. More than that, however, it systematizes searching for the corresponding formalism in other contexts. In fact, Marsden, Patrick and Shkoller [1998] show how this works in the context of classical field theory and multisymplectic geometry; *i.e.*, for Lagrangian and Hamiltonian pde's such as nonlinear wave equations.

Veselov discretizations of mechanics. We now show how the discrete Lagrangian formalism in Veselov [1988], [1991] and Moser and Veselov [1991] described earlier fits into this variational framework. Recall that a **discrete Lagrangian** is a smooth map $\mathbb{L} : Q \times Q \rightarrow \mathbb{R}$, and the corresponding action is

$$\mathbb{S} \equiv \sum_{k=0}^{n-1} \mathbb{L}(q_{k+1}, q_k).$$

The **discrete variational principle** is to extremize \mathfrak{S} for variations holding the endpoints q_0 and q_n fixed. This variational principle determines a “discrete flow” $\Phi : Q \times Q \rightarrow Q \times Q$ by $F(q_1, q_0) = (q_2, q_1)$, where q_2 is found from the **discrete Euler-Lagrange equations** (DEL equations):

$$\frac{\partial \mathbb{L}}{\partial q_1}(q_1, q_0) + \frac{\partial \mathbb{L}}{\partial q_0}(q_2, q_1) = 0. \quad (5.27)$$

In this section we work out the basic differential-geometric objects of this discrete mechanics *directly from the variational point of view*.

The Lagrange 1-form. We begin by calculating $d\mathfrak{S}$ for variations that do not fix the endpoints:

$$\begin{aligned} & d\mathfrak{S}(q_0, \dots, q_n) \cdot (\delta q_0, \dots, \delta q_n) \\ &= \sum_{k=0}^{n-1} \left(\frac{\partial \mathbb{L}}{\partial q_1}(q_{k+1}, q_k) \delta q_{k+1} + \frac{\partial \mathbb{L}}{\partial q_0}(q_{k+1}, q_k) \delta q_k \right) \\ &= \sum_{k=1}^n \frac{\partial \mathbb{L}}{\partial q_1}(q_k, q_{k-1}) \delta q_k + \sum_{k=0}^{n-1} \frac{\partial \mathbb{L}}{\partial q_0}(q_{k+1}, q_k) \delta q_k \\ &= \sum_{k=1}^n \left(\frac{\partial \mathbb{L}}{\partial q_1}(q_k, q_{k-1}) + \frac{\partial \mathbb{L}}{\partial q_0}(q_{k+1}, q_k) \right) \delta q_k \\ &\quad + \frac{\partial \mathbb{L}}{\partial q_0}(q_1, q_0) \delta q_0 + \frac{\partial \mathbb{L}}{\partial q_1}(q_n, q_{n-1}) \delta q_n. \end{aligned} \quad (5.28)$$

It is the last two terms that arise from the boundary variations (i.e. these are the ones that are zero if the boundary is fixed), and so these are the terms amongst which we expect to find the discrete analogue of the Lagrange 1-form. In fact, the boundary terms gives the *two* 1-forms on $Q \times Q$

$$\Theta_L^-(q_1, q_0) \cdot (\delta q_1, \delta q_0) \equiv \frac{\partial \mathbb{L}}{\partial q_0}(q_1, q_0) \delta q_0, \quad (5.29)$$

and

$$\Theta_L^+(q_1, q_0) \cdot (\delta q_1, \delta q_0) \equiv \frac{\partial \mathbb{L}}{\partial q_1}(q_1, q_0) \delta q_1, \quad (5.30)$$

and we regard *the pair* (Θ^-, Θ^+) as being the analogue of the 1-form in this situation.

Symplecticity of the flow. We parametrize the solutions of the variational principle by the initial conditions (q_1, q_0) , and restrict \mathfrak{S} to that solution space. Then equation (5.28) becomes

$$d\mathfrak{S} = \Theta_L^- + \Phi^* \Theta_L^+. \quad (5.31)$$

We should be able to obtain the symplecticity of Φ by determining what the equation $dd\mathfrak{S} = 0$ means for the right-hand-side of (5.31). At first, this does not appear to work, since $dd\mathfrak{S} = 0$ gives

$$\Phi^*(d\Theta_L^+) = -d\Theta_L^-, \quad (5.32)$$

which apparently says that Φ pulls a certain 2-form back to a different 2-form. The situation is aided by the observation that, from (5.29) and (5.30),

$$\Theta_L^- + \Theta_L^+ = \mathbf{d}\mathbb{L}, \quad (5.33)$$

and consequently,

$$\mathbf{d}\Theta_L^- + \mathbf{d}\Theta_L^+ = 0.$$

So there are *two* generally distinct 1-forms, but (up to sign) only *one* 2-form. If we make the definition

$$\Omega_L \equiv \mathbf{d}\Theta_L^- = -\mathbf{d}\Theta_L^+,$$

then (5.32) becomes $\Phi^*\Omega_L = \Omega_L$. Equation (5.29), in coordinates, gives

$$\Omega_L = \frac{\partial^2 \mathbb{L}}{\partial q_0^i \partial q_1^j} dq_0^i \wedge dq_1^j,$$

which agrees the discrete symplectic form discussed earlier.

Noether's Theorem. Suppose a Lie group G with Lie algebra \mathfrak{g} acts on Q , and hence diagonally on $Q \times Q$, and that \mathbb{L} is G -invariant. Clearly, \mathbb{S} is also G -invariant and G sends critical points of \mathbb{S} to themselves. Thus, the action of G restricts to the space of solutions, the map Φ is G -equivariant, and from (5.31),

$$0 = \xi_{Q \times Q} \lrcorner \mathbf{d}\mathbb{S} = \xi_{Q \times Q} \lrcorner \Theta_L^- + \xi_{Q \times Q} \lrcorner (\Phi^* \Theta_L^+),$$

for all $\xi \in \mathfrak{g}$, or equivalently, using equivariance of Φ ,

$$\xi_{Q \times Q} \lrcorner \Theta_L^- = -\Phi^*(\xi_{Q \times Q} \lrcorner \Theta_L^+). \quad (5.34)$$

Since \mathbb{L} is G -invariant, (5.33) gives $\xi_{Q \times Q} \lrcorner \Theta_L^- = -\xi_{Q \times Q} \lrcorner \Theta_L^+$, which in turn converts (5.34) to the conservation equation

$$\xi_{Q \times Q} \lrcorner \Theta_L^+ = \Phi^*(\xi_{Q \times Q} \lrcorner \Theta_L^+). \quad (5.35)$$

Defining the discrete momentum to be

$$J_\xi \equiv \xi_{Q \times Q} \lrcorner \Theta_L^+,$$

we see that (5.35) becomes conservation of momentum, recovering the conservation property of these integrators we found earlier. Marsden, Patrick and Shkoller [1998] develop a similar program for pde's in the context of multisymplectic geometry. Following this work, we take some first steps in this direction.

Multisymplectic geometry. We recall some aspects of multisymplectic geometry, following Gotay, Isenberg and Marsden [1997] and Marsden and Shkoller [1997].

We let $\pi_{XY} : Y \rightarrow X$ be a fiber bundle over an oriented manifold X . Denote the first jet bundle over Y by $J^1(Y)$ and identify it with the *affine* bundle over Y whose fiber over $y \in Y_x := \pi_{XY}^{-1}(x)$ consists of those linear mappings $\gamma : T_x X \rightarrow T_y Y$ satisfying

$$T\pi_{XY} \circ \gamma = \text{Identity on } T_x X.$$

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

$$x^\mu \mapsto (x^\mu, \phi^A(x^\mu), \partial_\nu \phi^A(x^\mu)), \quad (5.36)$$

where $\partial_\nu = \partial/\partial x^\nu$.

Definition 5.6 The **dual jet bundle** $J^1(Y)^*$ is the vector bundle over Y whose fiber at $y \in Y_x$ is the set of affine maps from $J^1(Y)_y$ to $\Lambda^{n+1}(X)_x$, the bundle of $(n + 1)$ -forms on X . A smooth section of $J^1(Y)^*$ is therefore an affine bundle map of $J^1(Y)$ to $\Lambda^{n+1}(X)$ covering π_{XY} .

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

$$v^A_\mu \mapsto (p + p_A^\mu v^A_\mu) d^{n+1}x, \quad (5.37)$$

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

Analogous to the canonical one- and two-forms on a cotangent bundle, there are canonical $(n + 1)$ - and $(n + 2)$ -forms on the dual jet bundle $J^1(Y)^*$. In coordinates, these forms are given by

$$\Theta = p_A^\mu dy^A \wedge d^n x_\mu + p d^{n+1}x \quad (5.38)$$

and

$$\Omega = dy^A \wedge dp_A^\mu \wedge d^n x_\mu - dp \wedge d^{n+1}x. \quad (5.39)$$

A **Lagrangian density** $\mathcal{L} : J^1(Y) \rightarrow \Lambda^{n+1}(X)$ is a smooth bundle map over X . In coordinates, we write

$$\mathcal{L}(\gamma) = L(x^\mu, y^A, v^A_\mu) d^{n+1}x. \quad (5.40)$$

The corresponding covariant Legendre transformation for \mathcal{L} is a fiber preserving map over Y , $\mathbb{F}\mathcal{L} : J^1(Y) \rightarrow J^1(Y)^*$, expressed intrinsically as the first order vertical Taylor approximation to \mathcal{L} :

$$\mathbb{F}\mathcal{L}(\gamma) \cdot \gamma' = \mathcal{L}(\gamma) + \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{L}(\gamma + \varepsilon(\gamma' - \gamma)) \quad (5.41)$$

where $\gamma, \gamma' \in J^1(Y)_y$. A straightforward calculation shows that the covariant Legendre transformation is given in coordinates by

$$p_A^\mu = \frac{\partial L}{\partial v^A_\mu}, \quad \text{and} \quad p = L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu. \quad (5.42)$$

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

$$\Theta_{\mathcal{L}} = (\mathbb{F}\mathcal{L})^* \Theta, \quad (5.43)$$

and the $(n + 2)$ -form $\Omega_{\mathcal{L}}$ by

$$\Omega_{\mathcal{L}} = -\mathbf{d}\Theta_{\mathcal{L}} = (\mathbb{R}\mathcal{L})^*\Omega, \quad (5.44)$$

with local coordinate expressions

$$\begin{aligned} \Theta_{\mathcal{L}} &= \frac{\partial L}{\partial v^A} dy^A \wedge d^n x_\mu + \left(L - \frac{\partial L}{\partial v^A} v^A \right) d^{n+1} x, \\ \Omega_{\mathcal{L}} &= dy^A \wedge d \left(\frac{\partial L}{\partial v^A} \right) \wedge d^n x_\mu - d \left[L - \frac{\partial L}{\partial v^A} v^A \right] \wedge d^{n+1} x. \end{aligned} \quad (5.45)$$

This formalism can be used to express, in an intrinsic way, the Euler–Lagrange equations, which in coordinates take the standard form

$$\frac{\partial L}{\partial y^A} (j^1(\phi)) - \frac{\partial}{\partial x^\mu} \left(\frac{\partial L}{\partial v^A} (j^1(\phi)) \right) = 0 \quad (5.46)$$

for a (local) section ϕ of Y . The left hand side of (5.46) is often denoted $\delta L/\delta\phi^A$ and is called the **Euler–Lagrange derivative** of L .

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

1. ϕ is a stationary point of $\int_X \mathcal{L}(j^1(\phi))$;
2. the Euler–Lagrange equations (5.46) hold in coordinates;
3. for any vector field W on J^1Y ,

$$(j^1(\phi))^* (\mathbf{i}_W \Omega_{\mathcal{L}}) = 0. \quad (5.47)$$

Example: A nonlinear wave equation. We consider a simple example to illustrate what the multisymplectic formalism is about. consider the nonlinear wave equation

$$\frac{\partial^2 \phi}{\partial t^2} - \Delta \phi - N'(\phi) = 0,$$

where Δ is the Laplace operator and N is a real-valued C^∞ function of one variable. Specialize now to one spatial dimension (the Laplace operator is then just the second spatial partial derivative operator) so that $n = 1$, $X = \mathbb{R}^2$, and the fibers of Y are \mathbb{R} . The Lagrangian density for this equation is

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

In coordinates $(x, t, \phi, \phi_t, \phi_x)$ on $J^1(Y)$, the multisymplectic 3-form is:

$$\begin{aligned} \Omega_{\mathcal{L}} &= -d\phi \wedge d\phi_t \wedge dx - d\phi \wedge d\phi_x \wedge dt \\ &\quad - N'(\phi) d\phi \wedge dx \wedge dt \\ &\quad + \phi_t d\phi_t \wedge dx \wedge dt - \phi_x d\phi_x \wedge dx \wedge dt; \end{aligned}$$

Sections of Y are mappings $(\phi(x, t))$ of \mathbb{R}^2 into \mathbb{R} , and sections of $J(Y)$ are mappings from \mathbb{R}^2 to \mathbb{R}^3 . Concretely, the first jet of a section ϕ is $j^1(\phi)(x, t) := (\phi(x, t), \phi_t(x, t), \phi_x(x, t))$. Write the conjugate momenta as $p^1 = \phi_x$ and $p^0 = \phi_t$.

The equation can be reformulated as

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ p^0 \\ p^1 \end{bmatrix}_{,0} + \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ p^0 \\ p^1 \end{bmatrix}_{,1} = \begin{bmatrix} N'(\phi) \\ -p^0 \\ p^1 \end{bmatrix}. \quad (5.48)$$

To each coefficient matrix \mathbf{J}_μ , we associate the contact form ω^μ on \mathbb{R}^3 given by $\omega^\mu(u_1, u_2) = \langle \mathbf{J}_\mu u_1, u_2 \rangle$, where $u_1, u_2 \in \mathbb{R}^3$ and $\langle \cdot, \cdot \rangle$ is the standard inner product on \mathbb{R}^3 . Thus, one sees in this example, the origin of the term “multisymplectic.”

Gotay, Isenberg and Marsden [1998] and Marsden, Patrick and Shkoller [1998] investigate the general theory of multisymplectic systems. One such interesting question is the PDE version of symplecticity of the flow, which is a general identity that the linearized equations must satisfy. In this case, this reads:

$$\frac{\partial}{\partial t} [\omega^0(j^1(\phi_t), j^1(\phi_x))] + \frac{\partial}{\partial x} [\omega^1(j^1(\phi_t), j^1(\phi_x))] = 0.$$

Many basic equations in continuum mechanics (wave equations, shallow water equations, etc.) have been put into the multisymplectic formalism (see Marsden and Shkoller [1997] for some simple examples and for links with the important work of Bridges [1994, 1997]). Marsden, Patrick and Shkoller [1998] show how to obtain all of the multisymplectic and Noether theorem structure directly from the variational principle we have described above entirely on the Lagrangian side, $J^1(Y)$.

This is then the beginnings of the development of the basic differential-geometric formulation of the multisymplectic structure that can be used for a variety of purposes, including Hamiltonian and Lagrangian PDE's. Using this framework, Marsden, Patrick and Shkoller [1998] produce multisymplectic variational integrators of Veselov type and test these *multisymplectic integrators* numerically. For example, in a test involving 5000 collisions of kink solitons (see Figure 5.2) for the sine-Gordon equation on a circle, the method does very well indeed.

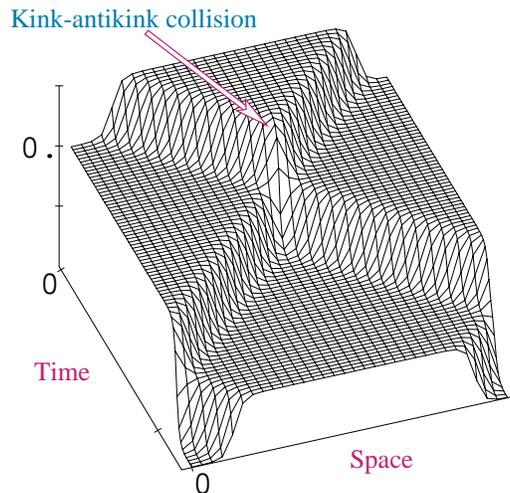


Figure 5.2: Collision of a kink-antikink pair for the sine-Gordon equation.

References

- Abarbanel, H.D.I. and D.D. Holm [1987] Nonlinear stability analysis of inviscid flows in three dimensions: incompressible fluids and barotropic fluids, *Phys. Fluids* **30**, 3369–3382.
- Abarbanel, H.D.I., D.D. Holm, J.E. Marsden, and T.S. Ratiu [1986] Nonlinear stability analysis of stratified fluid equilibria, *Phil. Trans. Roy. Soc. London A* **318**, 349–409; also Richardson number criterion for the nonlinear stability of three-dimensional stratified flow, *Phys. Rev. Lett.* **52** [1984], 2552–2555.
- Abraham, R. and J.E. Marsden [1978] *Foundations of Mechanics*, Second Edition, Addison-Wesley.
- Abraham, R., J.E. Marsden, and T.S. Ratiu [1988] *Manifolds, Tensor Analysis, and Applications*. Second Edition, Applied Mathematical Sciences **75**, Springer-Verlag.
- Adams, M. and T.S. Ratiu [1988] The three point vortex problem: commutative and non-commutative integrability, in *Hamiltonian Dynamical Systems* (K. Meyer, D. Saari, eds.), *Cont. Math.* **81**, 245–257.
- Alber, M.S., R. Camassa, D.D. Holm and J.E. Marsden [1994], The geometry of peaked solitons and billiard solutions of a class of integrable pde's, *Lett. Math. Phys.* **32**, 137–151.
- Alber, M.S., R. Camassa, D.D. Holm and J.E. Marsden [1995], On the link between umbilic geodesics and soliton solutions of nonlinear PDE's, *Proc. Roy. Soc* **450**, 677–692.
- Alber, M.S., G.G. Luther, J.E. Marsden and J.W. Robbins [1998a] Geometry and control of $\chi^{(2)}$ processes and the generalized Poincaré sphere, *preprint*.
- Alber, M.S., G.G. Luther, J.E. Marsden and J.W. Robbins [1998b], Geometric phases, reduction and Lie-Poisson structure for the resonant three-wave interaction, *preprint*.
- Armero, F. and J.C. Simo [1992] A New Unconditionally Stable Fractional Step Method for Non-Linear Coupled Thermomechanical Problems. *Int. J. Num. Meth. Eng.* **35**, 737–766.
- Armero, F. and J.C. Simo [1993] A-Priori Stability Estimates and Unconditionally Stable Product Formula Algorithms for Non-Linear Coupled Thermoplasticity. *Int. J of Plasticity*, **9**, 149–182.
- Armero, F. and J.C. Simo [1996] Long-Term Dissipativity of Time-Stepping Algorithms for an Abstract Evolution Equation with Applications to the Incompressible MHD and Navier-Stokes Equations. *Comp. Meth. Appl. Mech. Eng.* **131**, 41–90.
- Armero, F. and J.C. Simo [1996] Formulation of a new class of fraction-step methods for the incompressible MHD equations that retains the long-term dissipativity of the continuum dynamical system, *Fields Institute Comm.* **10**, 1–23.
- Arms, J.M., J.E. Marsden, and V. Moncrief [1981] Symmetry and bifurcations of momentum mappings, *Comm. Math. Phys.* **78**, 455–478.
- Arms, J.M., J.E. Marsden, and V. Moncrief [1982] The structure of the space solutions of Einstein's equations: II Several Killings fields and the Einstein-Yang-Mills equations, *Ann. of Phys.* **144**, 81–106.

- Arnold, V.I. [1966a] Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier, Grenoble* **16**, 319–361.
- Arnold, V.I. [1966b] On an a priori estimate in the theory of hydrodynamical stability. *Izv. Vyssh. Uchebn. Zaved. Mat. Nauk* **54**, 3–5; English Translation: *Amer. Math. Soc. Transl.* **79** [1969], 267–269.
- Arnold, V.I. [1966c] Sur un principe variationnel pour les découlements stationnaires des liquides parfaits et ses applications aux problèmes de stabilité non linéaires. *J. Mécanique* **5**, 29–43.
- Arnold, V. I. [1989] *Mathematical Methods of Classical Mechanics*. Second Edition Graduate Texts in Math **60**, Springer-Verlag.
- Arnold, V.I., V.V. Kozlov, and A.I. Neishtadt [1988] Mathematical aspects of classical and celestial mechanics, in: *Dynamical Systems III*, V.I. Arnold, ed. Springer-Verlag.
- Arnold, V.I. and B. Khesin [1992] Topological methods in hydrodynamics. *Ann. Rev. Fluid Mech.* **24**, 145–166.
- Arnold, V.I. and B. Khesin [1997] *Topological methods in Fluid Dynamics*. Appl. Math. Sciences, Springer-Verlag.
- Ashbaugh, M.S., C.C. Chicone, and R.H. Cushman [1990] The twisting tennis racket. *Dyn. Diff. Eqns.* **3**, 67–85.
- Åström, K.J. and K. Furuta [1996] Swinging up a pendulum by energy control. *IFAC, San Francisco* **13**.
- Austin, M., P.S. Krishnaprasad, and L.S. Wang [1993] Almost Poisson integration of rigid body systems. *J. Comput. Phys.* **107**, 105–117.
- Barth, E. and Leimkuhler, B. [1996a]. A semi-explicit, variable-stepsize integrator for constrained dynamics. Mathematics department preprint series, University of Kansas.
- Barth, E. and Leimkuhler, B. [1996b]. Symplectic methods for conservative multibody systems. *Fields Institute Communications*, **10**, 25–43.
- Bates, L. and R. Cushman [1997] *Global Aspects of Classical Integrable Systems*, Birkhäuser, Boston.
- Bates, L. and E. Lerman [1996] Singular reduction at nonzero momentum values *Pacific J. Math.* (to appear).
- Bates, L. and J. Sniatycki [1993] Nonholonomic reduction, *Reports on Math. Phys.* **32**, 99–115.
- Benjamin, T.B. [1972] The stability of solitary waves. *Proc. Roy. Soc. A.* **328**, 153–183.
- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and R. Murray [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.*, **136**, 21–99.
- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1994] Dissipation Induced Instabilities, *Ann. Inst. H. Poincaré, Analyse Nonlineaire* **11**, 37–90.
- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1996] The Euler-Poincaré equations and double bracket dissipation. *Comm. Math. Phys.* **175**, 1–42.

- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and G. Sánchez de Alvarez [1992] Stabilization of rigid body dynamics by internal and external torques, *Automatica* **28**, 745–756.
- Bloch, A.M., N. Leonard and J.E. Marsden [1997] Stabilization of Mechanical Systems Using Controlled Lagrangians, *CDC, IEEE, San Diego*.
- Bloch, A.M., N. Leonard and J.E. Marsden [1998] Controlled Lagrangians and the Stabilization of Mechanical Systems, *in preparation*.
- Bloch, A.M. and J.E. Marsden [1989] Controlling homoclinic orbits, *Theor. and Comp. Fluid Mech.* **1**, 179–190.
- Bloch, A.M. and J.E. Marsden [1990] Stabilization of rigid body dynamics by the energy-Casimir method, *Systems Control Lett.* **14**, 341–346.
- Bloch, A.M., J.E. Marsden and G. Sánchez [1997] Stabilization of relative equilibria of mechanical systems with symmetry, *Current and Future Directions in Applied Mathematics*, Edited by M. Alber, B. Hu, and J. Rosenthal, Birkhäuser, 43–64.
- Bobenko, A.I., A.G. Reyman and M.A. Semenov-Tian-Shansky [1989] The Kowalewski Top 99 years later: A Lax pair, generalizations and explicit solutions, *Comm. Math. Phys.* **122**, 321–354.
- Bona, J. [1975] On the stability theory of solitary waves. *Proc. Roy. Soc. London* **344A**, 363–374.
- Bretherton, F.P. [1970] A note on Hamilton’s principle for perfect fluids. *J. Fluid Mech.* **44**, 19–31.
- Bridges, T.J. [1994] *Hamiltonian spatial structure for 3D water waves relative to a moving frame of reference*, *J. Nonlinear Sci.* **4**, 221–251.
- Bridges, T.J. [1997] Multi-symplectic structures and wave propagation, *Math Proc Camb Phil Soc*, **121**, 147–190.
- Brockett, R.W. [1981] Control theory and singular Riemannian geometry, in *New Directions in Applied Mathematics*, P.J. Hilton and G.S. Young (eds.), Springer-Verlag.
- Camassa, R. and D.D. Holm [1993] An integrable shallow water equation with peaked solitons, *Phys. Rev. Lett.*, **71**, 1661–1664.
- Camassa, R., D.D. Holm and J.M. Hyman [1994] A new integrable shallow water equation, *Adv. Appl. Mech.*, **31**, 1–33.
- Cendra, H., D.D. Holm, M.J.W. Hoyle and J. E. Marsden [1998] The Maxwell-Vlasov equations in Euler-Poincaré form. *J. Math. Phys (to appear)*.
- Cendra, H., D.D. Holm, J. E. Marsden and T.S. Ratiu [1998] Lagrangian Reduction, the Euler-Poincaré Equations, and Semidirect Products. *Arnold volume II, AMS (to appear)*.
- Cendra, H., A. Ibrort, and J.E. Marsden [1987] Variational principal fiber bundles: a geometric theory of Clebsch potentials and Lin constraints, *J. Geom. Phys.* **4**, 183–206.

- Cendra, H. and J.E. Marsden [1987] Lin constraints, Clebsch potentials and variational principles, *Physica D* **27**, 63–89.
- Cendra, H., J. E. Marsden and T.S. Ratiu [1998] Lagrangian reduction by stages. *preprint*.
- Channell, P. and C. Scovel [1990] Symplectic integration of Hamiltonian systems, *Nonlinearity* **3**, 231–259.
- Channell, P. and C. Scovel [1991] Integrators for Lie-Poisson Dynamical Systems, *Physica D*, **50**, 80–88.
- Chern, S.J. and J.E. Marsden [1990] A note on symmetry and stability for fluid flows, *Geo. Astro. Fluid. Dyn.* **51**, 1–4.
- Chetayev, N.G. [1941] On the equations of Poincaré. *J. Appl. Math. Mech.* **5**, 253–262
- Chandrasekhar, S. [1967] Ellipsoidal figures of equilibrium-an historical account. *Comm. on Pure and Math* **20**, 251–265.
- Chandrasekhar, K. [1977] Ellipsoidal Figures of Equilibrium. Dover.
- Chorin, A.J, T.J.R. Hughes, J.E. Marsden, and M. McCracken [1978] Product Formulas and Numerical Algorithms, *Comm. Pure Appl. Math.* **31**, 205–256.
- Chorin, A.J. and J.E. Marsden [1989] *A Mathematical Introduction to Fluid Mechanics*. Texts in Applied Mathematics, **4**, Springer-Verlag, New York.
- Churchill, R.C., M. Kummer, and D.L. Rod [1983] On averaging, reduction and symmetry in Hamiltonian systems. *J. Diff. Eqns.* **49**, 359–373.
- Clebsch, A. [1857] Über eine allgemeine Transformation der hydrodynamischen Gleichungen, *Z. Reine Angew. Math.* **54**, 293–312.
- Clebsch, A. [1859] Über die Integration der hydrodynamischen Gleichungen, *Z. Reine Angew. Math.* **56**, 1–10.
- Cushman, R., J. Hermans, and D. Kemppainen [1995] The rolling disc. in *Nonlinear dynamical systems and chaos (Groningen, 1995)*, Progr. Nonlinear Differential Equations Appl., **19**, Birkhduser, Basel, 21–60.
- Cushman, R., D. Kemppainen, J. Śniatycki, and L. Bates [1995] Geometry of nonholonomic constraints. *Rep. Math. Phys.* **36**, 275–286.
- Cushman, R. and D. Rod [1982] Reduction of the semi-simple 1:1 resonance, *Physica D* **6**, 105–112.
- Cushman, R. and R. Sjamaar [1991] On singular reduction of Hamiltonian spaces, *Symplectic Geometry and Mathematical Physics*, ed. by P. Donato, C. Duval, J. Elhadad, and G.M. Tuynman, Birkhäuser, 114–128.
- David, D., Holm, D. and Tratnik, M. V. [1989] Integrable and chaotic polarization dynamics in nonlinear optical beams. *Physics Lett. A* **137**, 355–364.
- David, D., D.D. Holm, and M. Tratnik [1990] Hamiltonian chaos in nonlinear optical polarization dynamics, *Phys. Rep.* **187**, 281–370.
- David, D., D.D. Holm [1990] Multiple Lie-Poisson structures, reductions, and geometric phases for the Maxwell-Bloch travelling wave equations, *J. Nonlinear Sci.* **2**, 241–262.

- Dellnitz, M., J.E. Marsden, I. Melbourne, and J. Scheurle [1992] Generic bifurcations of pendula. *Int. Series on Num. Math.* **104**, 111-122. ed. by G. Allgower, K. Böhmer, and M. Golubitsky, Birkhäuser.
- Dellnitz, M. and I. Melbourne [1993] The equivariant Darboux theorem. *Lect. Appl. Math.* **29**, 163–169.
- Dellnitz, M., I. Melbourne, and J.E. Marsden [1992] Generic bifurcation of Hamiltonian vector fields with symmetry, *Nonlinearity* **5**, 979–996.
- Ebin, D.G. and J.E. Marsden [1970] Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. Math.* **92**, 102–163.
- Feng, K. [1986] Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comp. Math.* **4**, 279–289.
- Feng, K. and Z. Ge [1988] On approximations of Hamiltonian systems, *J. Comp. Math.* **6**, 88–97.
- Feng, K. and M.Z. Qin [1987] The symplectic methods for the computation of Hamiltonian equations, *Springer Lecture Notes in Math.* **1297**, 1–37.
- Ge, Z., H.P. Kruse and J.E. Marsden [1996] The limits of Hamiltonian structures in three-dimensional elasticity, shells and rods *J. Nonlin. Sci.* **6**, 19–57.
- Ge, Z., H.P. Kruse, J.E. Marsden and C. Scovel [1995] The Convergence of Hamiltonian Structures in the Shallow Water Approximation (*Canadian Quart. of Appl. Math.*), **3**, 277–302.
- Ge, Z. and J.E. Marsden [1988] Lie-Poisson integrators and Lie-Poisson Hamilton-Jacobi theory, *Phys. Lett. A* **133**, 134–139.
- Getz, N.H. and J. E. Marsden [1995] Control for an autonomous bicycle, *International Conference on Robotics and Automation, IEEE*, Nagoya, Japan, May, 1995.
- Gillilan, R. and K. Wilson [1992]. Shadowing, rare events, and rubber bands: A variational Verlet algorithm for molecular dynamics. *J. Chem. Phys.*, **97**, 1757–1772.
- Gjaja, I. and D.D. Holm [1996] Self-consistent wave-mean flow interaction dynamics and its Hamiltonian formulation for a rotating stratified incompressible fluid, *Physica D*, **98** (1996) 343-378.
- Golubitsky, M., I. Stewart, and D. Schaeffer [1988] *Singularities and Groups in Bifurcation Theory*. Vol. 2, Applied Mathematical Sciences **69**, Springer-Verlag.
- Gonzalez, O. and J. C. Simo [1995] On the Stability of Symplectic and Energy-Momentum Algorithms for Nonlinear Hamiltonian Systems with Symmetry. Submitted to *Comp. Meth. Appl. Mech. Eng.*.
- Gonzalez, O. [1996] *Design and Analysis of Conserving Integrators for Nonlinear Hamiltonian Systems with Symmetry*, Thesis, Stanford University, Mechanical Engineering.
- Gonzalez, O. [1996]. Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.* **6**, 449–468.
- Gotay, M., J. Isenberg, and J.E. Marsden [1997] *Momentum Maps and the Hamiltonian Structure of Classical Relativistic Field Theories, I*. preprint.

- Greenspan, D. [1974] *Discrete Numerical Methods in Physics and Engineering*. Academic Press.
- Greenspan, D. [1984] Conservative numerical methods for $\ddot{x} = f(x)$, *J. Comp. Phys.* **56**, 21–48.
- Grillakis, M., J. Shatah and W. Strauss [1987] Stability theory of solitary waves in the presence of symmetry. I & II. *J. Funct. Anal.* **74**, 160–197 and **94** (1990), 308–348.
- Gozzi, E. and W.D. Thacker [1987] Classical adiabatic holonomy in a Grassmannian system. *Phys. Rev. D* **35**, 2388–2396.
- Guckenheimer, J. and A. Mahalov [1992], Resonant triad interactions in symmetric systems, *Physica D* **54**, 267–310.
- Guichardet, A. [1984] On rotation and vibration motions of molecules, *Ann. Inst. H. Poincaré* **40**, 329–342.
- Guillemin, V., Lerman, E. and Sternberg, S. [1996] *Symplectic Fibrations and Multiplicity Diagrams*. LMS Lecture Note Series, Cambridge University Press.
- Guillemin, V. and Sternberg, S. [1978] On the equations of motions of a classic particle in a Yang-Mills field and the principle of general covariance. *Hadronic J.* **1**, 1–32.
- Guillemin, V. and S. Sternberg [1980] The moment map and collective motion, *Ann. of Phys.* **1278**, 220–253.
- Guillemin, V. and S. Sternberg [1982] Convexity properties of the moment map. *Invent. Math.* **67**, 491–513, **77**, 533–546.
- Guillemin, V. and S. Sternberg [1984] *Symplectic Techniques in Physics*. Cambridge University Press.
- Guo, Y. and Strauss, W.A. [1995a] Instability of periodic BGK equilibria. *Comm. Pure Appl. Math.* **48**, 861–894.
- Guo, Y. and Strauss, W.A. [1995b] Nonlinear instability of double-humped equilibria. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **12**, 339–352.
- Hamel, G. [1904] Die Lagrange-Eulerschen Gleichungen der Mechanik. *Z. für Mathematik u. Physik* **50**, 1–57.
- Hamel, G. [1949] *Theoretische Mechanik*, Springer-Verlag.
- Holm, D.D. [1987] Hamiltonian dynamics and stability analysis of neutral electromagnetic fluids with induction, *Physica D* **25** (1987) 261–287.
- Holm, D.D. [1996] Hamiltonian balance equations, *Physica D*, **98** (1996) 379–414.
- Holm, D.D., S. Kouranbaeva, J.E. Marsden, T. Ratiu and S. Shkoller [1998], Euler-Poincaré equations for continuum theories on Riemannian manifolds. *In preparation*.
- Holm, D.D. and B.A. Kupershmidt [1983] Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity. *Physica D* **6**, 347–363.
- Holm, D.D., J.E. Marsden, and T.S. Ratiu [1986] The Hamiltonian structure of continuum mechanics in material, spatial and convective representations, *Séminaire de Mathématiques supérieures, Les Presses de L’Univ. de Montréal* **100**, 11–122.

- Holm, D. D., Marsden, J. E. and Ratiu, T. [1998a] The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Adv. in Math.* (to appear).
- Holm, D. D., J.E. Marsden and T. Ratiu [1998b] The Euler-Poincaré equations in geophysical fluid dynamics, in *Proceedings of the Isaac Newton Institute Programme on the Mathematics of Atmospheric and Ocean Dynamics*, Cambridge University Press (to appear).
- Holm, D.D., J.E. Marsden, T.S. Ratiu, and A. Weinstein [1985] Nonlinear stability of fluid and plasma equilibria, *Phys. Rep.* **123**, 1–116.
- Holm, D.D. and V. Zeitlin [1997] Hamilton’s principle for quasigeostrophic motion, *In preparation*.
- Holmes, P., J. R. Jenkins, and N. Leonard [1997] Dynamics of the Kirchhoff equations I: coincident centers of gravity and buoyancy, (preprint, Princeton University.)
- Holmes, P.J. and J.E. Marsden [1978] Bifurcation to divergence and flutter in flow-induced oscillations: an infinite dimensional analysis, *Automatica*, **14**, 367–384.
- Holmes, P.J. and J.E. Marsden [1981] A partial differential equation with infinitely many periodic orbits: chaotic oscillations of a forced beam, *Arch. Rat. Mech. Anal.* **76**, 135–166.
- Holmes, P.J. and J.E. Marsden [1982a] Horseshoes in perturbations of Hamiltonian systems with two degrees of freedom, *Comm. Math. Phys.* **82**, 523–544.
- Holmes, P.J. and J.E. Marsden [1982b] Melnikov’s method and Arnold diffusion for perturbations of integrable Hamiltonian systems, *J. Math. Phys.* **23**, 669–675.
- Holmes, P.J. and J.E. Marsden [1983] Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups, *Indiana Univ. Math. J.* **32**, 273–310.
- Holmes, P.J., J.E. Marsden, and J. Scheurle [1988] Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations, *Cont. Math.* **81**, 213–244.
- Iwai, T. [1987] A geometric setting for classical molecular dynamics, *Ann. Inst. Henri Poincaré, Phys. Th.* **47**, 199–219.
- Iwai, T. [1990] On the Guichardet/Berry connection, *Phys. Lett. A* **149**, 341–344.
- Kazhdan, D., B. Kostant, and S. Sternberg [1978] Hamiltonian group actions and dynamical systems of Calogero type, *Comm. Pure Appl. Math.* **31**, 481–508.
- Kirillov, A.A. [1962] Unitary representations of nilpotent Lie groups, *Russian Math. Surveys* **17**, 53–104.
- Kirillov, A.A. [1976a] Local Lie Algebras. *Russian Math. Surveys* **31**, 55–75.
- Kirillov, A.A. [1976b] *Elements of the Theory of Representations*. Grundlehren der math. Wiss. **220** Springer-Verlag.
- Kirk, V., J.E. Marsden, and M.Silber [1996] Branches of stable three-tori using Hamiltonian methods in Hopf bifurcation on a rhombic lattice, *Dyn. and Stab. of Systems*, **11**, 267–302.

- Knobloch, E., A. Mahalov, and J.E. Marsden [1994] Normal Forms for three-dimensional Parametric Instabilities in Ideal Hydrodynamics *Physica D* **73**, 49–81.
- Koon, W.S. and J.E. Marsden [1997] Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control and Optim.* **35**, 901–929.
- Koon, W.S. and J.E. Marsden [1998a] The Hamiltonian and Lagrangian Approaches to the Dynamics of Nonholonomic Systems, *Reports on Math Phys (to appear)*.
- Koon, W.S. and J.E. Marsden [1998b] The Geometric Structure of Nonholonomic Mechanics, *Proc. CDC, San Diego*, **FP12-4**, 4856–4862.
- Koon, W.S. and J.E. Marsden [1998c] The Poisson reduction of nonholonomic mechanical systems, *Reports on Math Phys (to appear)*.
- Korteweg, D. [1899] Ueber eine ziemlich verbreitete unrichtige Behandlungsweise eines Problemes der rollenden Bewegung und insbesondere über kleine rollende Schwingungen um eine Gleichgewichtslage. *Nieuw Archiefvoor Wiskunde.* **4**, 130–155.
- Korteweg, D.J. and G. de Vries [1895] On the change of form of long waves advancing in a rectangular canal and on a new type of long stationary wave. *Phil. Mag.* **39**, 422–433.
- Krishnaprasad, P.S. [1985] Lie-Poisson structures, dual-spin spacecraft and asymptotic stability, *Nonl. Anal. Th. Meth. and Appl.* **9**, 1011–1035.
- Krishnaprasad, P.S. and J.E. Marsden [1987] Hamiltonian structure and stability for rigid bodies with flexible attachments, *Arch. Rat. Mech. Anal.* **98**, 137–158.
- Krishnaprasad, P.S., J.E. Marsden, and T. Posbergh [1987] Stability analysis of a rigid body with a flexible attachment using the energy-Casimir method, *Cont. Math. AMS* **68**, 253–273.
- Kruse, H.P., J.E. Marsden, and J. Scheurle [1993] On uniformly rotating field drops trapped between two parallel plates. *Lect. in Appl. Math. AMS* **29**, 307–317.
- Kummer, M. [1981] On the construction of the reduced phase space of a Hamiltonian system with symmetry, *Indiana Univ. Math. J.* **30**, 281–291.
- Kummer, M. [1990] On resonant classical hamiltonians with n frequencies, *J. Diff. Eqns.* **83**, 220–243.
- Kupersmidt, B.A. and T. Ratiu [1983] Canonical maps between semidirect products with applications to elasticity and superfluids, *Comm. Math. Phys.* **90**, 235–250.
- Lamb, H. [1932] *Hydrodynamics*. Sixth Edition, Dover, NY.
- Lagrange, J.L. [1788] *Mécanique Analytique*. Chez la Veuve Desaint
- Leimkuhler, B. and G. Patrick [1996]. Symplectic integration on Riemannian manifolds. *J. of Nonl. Sci.* **6**, 367–384.
- Leimkuhler, B. and R. Skeel [1994]. Symplectic numerical integrators in constrained Hamiltonian systems. *Journal of Computational Physics*, **112**, 117–125.
- Leonard, N.E. [1995a] Control synthesis and adaptation for an underactuated autonomous underwater vehicle, *IEEE J. of Oceanic Eng.* **20**, 211–220.

- Leonard, N.E. [1995b] Periodic forcing, dynamics and control of underactuated spacecraft and underwater vehicles, *Proc. 34th IEEE Conf. Decision Ctrl.*, 3980–3985.
- Leonard, N.E. [1996] Stabilization of underwater vehicle dynamics with symmetry breaking potentials, *in preparation*. (See also Leonard, N.E., Stabilization of steady motions of an underwater vehicle, *Proc. 35th IEEE Conference on Decision and Control*, Kobe, Japan, 1996.)
- Leonard, N.E. [1997] Stability of a bottom-heavy underwater vehicle, *Automatica* **33**.
- Leonard, N.E. and P.S. Krishnaprasad [1995] Motion control of drift-free, left-invariant systems on Lie groups, *IEEE Trans. Automat. Control* **40**, 1539–1554.
- Leonard, N.E. and J.E. Marsden [1997] Stability and Drift of Underwater Vehicle Dynamics: Mechanical Systems with Rigid Motion Symmetry, *Physica D* **105**, 130–162.
- Lewis, A. [1996] Affine connections and distributions, *preprint*.
- Lewis, A. [1996] The geometry of the Gibbs-Appell equations and Gauss' principle of least constraint, *Reports on Math. Phys.* **38**, 11–28.
- Lewis, A. & R.M. Murray [1995] Variational principles in constrained systems: theory and experiments, *Int. J. Nonlinear Mech.* **30**, 793–815.
- Lewis, A., J.P. Ostrowski, R.M. Murray & J. Burdick [1994] Nonholonomic mechanics and locomotion: the snakeboard example. *IEEE Intern. Conf. on Robotics and Automation*.
- Lewis, D. [1989] Nonlinear stability of a rotating planar liquid drop, *Arch. Rat. Mech. Anal.* **106**, 287–333.
- Lewis, D. [1992a] Bifurcation of liquid drops. *Nonlinearity* **6**, 491–522.
- Lewis, D. [1992b] Lagrangian block diagonalization. *Dyn. Diff. Eqn's.* **4** 1–42.
- Lewis, D., J.E. Marsden, R. Montgomery, and T.S. Ratiu [1986] The Hamiltonian structure for dynamic free boundary problems, *Physica D* **18**, 391–404.
- Lewis, D., J.E. Marsden and T.S. Ratiu [1987] Stability and bifurcation of a rotating liquid drop, *J. Math. Phys.* **28**, 2508–2515.
- Lewis, D. and T.S. Ratiu [1996] Rotating n -gon/ kn -gon vortex configurations, *Journ. Nonlinear Science* **6**, 385–414.
- Lewis, D. and J.C. Simo [1990] Nonlinear stability of rotating pseudo-rigid bodies, *Proc. Roy. Soc. Lon. A* **427**, 281–319.
- Lewis, D., T.S. Ratiu, J.C. Simo, and J.E. Marsden [1992] The heavy top: a geometric treatment, *Nonlinearity* **5**, 1–48.
- Le, H. and D.G. Kendall [1993] The Riemannian structure of Euclidean shape spaces: a novel environment for statistics *Ann. of Statistics*, **21**, 1225–1271.
- Lie, S. [1890] *Theorie der Transformationsgruppen, Zweiter Abschnitt*. Teubner, Leipzig.
- Littlejohn, R. and M. Reinch [1997] Gauge fields in the separation of rotations and internal motions in the n -body problem, *Rev. Mod. Phys.* **69**, 213–275.

- Low, F.E. [1958] A Lagrangian formulation of the Boltzmann–Vlasov equation for plasmas, *Proc. Roy. Soc. Lond. A* **248**, 282–287.
- Maddocks, J. [1991] On the stability of relative equilibria. *IMA J. Appl. Math.* **46**, 71–99.
- Maddocks, J. and M. L. Overton [1995] Stability Theory for Dissipatively Perturbed Hamiltonian Systems, *Comm. Pure and Applied Math.* **XLVIII**, 583–610
- Maddocks, J. and R.L. Sachs [1993] On the stability of KdV multi-solitons, *Comm. Pure and Applied Math.* **46**, 867–901.
- Mahalov, A. and S. Nikitin [1997] Position controllability of a deformable body in ideal fluid. *Math. Models Methods Appl. Sci.* **7**, 139–149.
- Marsden, J.E. [1982] A group theoretic approach to the equations of plasma physics, *Can. Math. Bull.* **25**, 129–142.
- Marsden, J.E. [1992], *Lectures on Mechanics* London Mathematical Society Lecture note series, **174**, Cambridge University Press.
- Marsden, J.E., G. Misiolek, M. Perlmutter and T.S. Ratiu [1998a] Symplectic Reduction for Semidirect Products and Central Extensions, *Diff. Geometry and Appl.* (to appear).
- Marsden, J.E., G. Misiolek, M. Perlmutter and T.S. Ratiu [1998b] Reduction by stages and group extensions, *in preparation*.
- Marsden, J.E., R. Montgomery, and T.S. Ratiu [1990] *Reduction, symmetry, and phases in mechanics*. Memoirs AMS **436**.
- Marsden, J.E. and P.J. Morrison [1984] Noncanonical Hamiltonian field theory and reduced MHD. *Cont. Math. AMS*, **28**, 133–150.
- Marsden, J.E. and J. Ostrowski [1998] Symmetries in Motion: Geometric Foundations of Motion Control (to appear).
- Marsden, J.E., O.M. O’Reilly, F.J. Wicklin, and B.W. Zombro [1991] Symmetry, stability, geometric phases, and mechanical integrators, *Nonlinear Science Today* **1**, 4–11, **1**, 14–21.
- Marsden, J. E., G. W. Patrick, and W. F. Shadwick [1996] *Integration Algorithms and Classical Mechanics*. Fields Institute Communications, **10**, Am. Math. Society.
- Marsden, J.E., G.W. Patrick and S. Shkoller [1998] Variational methods in continuous and discrete mechanics and field theory. *preprint*.
- Marsden, J.E. and T.S. Ratiu [1986] Reduction of Poisson manifolds, *Lett. in Math. Phys.* **11**, 161–170.
- Marsden, J. E. and Ratiu, T.S., [1998] *Introduction to Mechanics and Symmetry* 2st Edition, Springer-Verlag, New York.
- Marsden, J.E., T.S. Ratiu, and G. Raugel [1995] Équations d’Euler dans une coque sphérique mince (Euler equations on a thin spherical shell), *C.R. Acad. Sci. Paris* **321**, Série I, 1201–1206.
- Marsden, J.E., T.S. Ratiu, and A. Weinstein [1984a] Semi-direct products and reduction in mechanics, *Trans. Am. Math. Soc.* **281**, 147–177.

- Marsden, J.E., T.S. Ratiu, and A. Weinstein [1984b] Reduction and Hamiltonian structures on duals of semidirect product Lie Algebras, *Cont. Math. AMS* **28**, 55–100.
- Marsden, J.E. and J. Scheurle [1993a] Lagrangian reduction and the double spherical pendulum, *ZAMP* **44**, 17–43.
- Marsden, J.E. and J. Scheurle [1993b] The reduced Euler-Lagrange equations, *Fields Institute Comm.* **1**, 139–164.
- Marsden, J.E. and J. Scheurle [1995] Pattern evocation and geometric phases in mechanical systems with symmetry, *Dyn. and Stab. of Systems.* **10**, 315–338.
- Marsden, J.E., J. Scheurle and J. Wendlandt [1996] Visualization of orbits and pattern evocation for the double spherical pendulum, *ICIAM 95: Mathematical Research, Akademie Verlag, Ed. by K. Kirchgässner, O. Mahrenholtz and R. Mennicken* **87**, 213–232.
- Marsden, J.E. and S. Shkoller [1997] Multisymplectic geometry, covariant Hamiltonians and water waves *Math. Proc. Camb. Phil. Soc.*, to appear.
- Marsden, J.E., and J.C. Simo [1990] The energy-momentum method, *La “Mécanique Analytique” de Lagrange et son Héritage, Atti della Accademia delle Scienze di Torino*, **124**, 245–268.
- Marsden, J.E., J.C. Simo, D.R. Lewis, and T.A. Posbergh [1989] A block diagonalization theorem in the energy momentum method, *Cont. Math. AMS* **97**, 297–313.
- Marsden, J.E. and A. Weinstein [1974] Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.* **5**, 121–130.
- Marsden, J.E. and A. Weinstein [1982] The Hamiltonian structure of the Maxwell-Vlasov equations, *Physica D* **4**, 394–406.
- Marsden, J.E. and A. Weinstein [1983] Coadjoint orbits, vortices and Clebsch variables for incompressible fluids, *Physica D* **7**, 305–323.
- Marsden, J.E., A. Weinstein, T.S. Ratiu, R. Schmid, and R.G. Spencer [1983] Hamiltonian systems with symmetry, coadjoint orbits and plasma physics, in Proc. IUTAM-IS1MM Symposium on *Modern Developments in Analytical Mechanics*, Torino 1982, *Atti della Acad. della Sc. di Torino* **117**, 289–340.
- Marsden, J.E. and J.M. Wendlandt [1997] Mechanical systems with symmetry, variational principles and integration algorithms. *Current and Future Directions in Applied Mathematics*, Edited by M. Alber, B. Hu, and J. Rosenthal, Birkhäuser, 219–261.
- McLachlan, R. I. and C. Scovel [1996]. A survey of open problems in symplectic integration. *Fields Institute Communications*, **10**, 151–180.
- McLaughlan, R.I. and C. Scovel [1995] Equivariant constrained symplectic integration. *J. of Nonl. Sci.* **5**, 233–256.
- Meyer, K.R. [1973] Symmetries and integrals in mechanics, in *Dynamical Systems*, M. Peixoto (ed.), Academic Press, 259–273.
- Misiolek, G. [1997] A shallow water equation as a geodesic flow on the Bott-Virasoro group. *J. Geom. Phys.*, to appear.

- Montgomery, R. [1984] Canonical formulations of a particle in a Yang-Mills field, *Lett. Math. Phys.* **8**, 59–67.
- Montgomery, R. [1986] The Bundle Picture in Mechanics, Ph.D. Thesis, Berkeley.
- Montgomery, R. [1988] The connection whose holonomy is the classical adiabatic angles of Hannay and Berry and its generalization to the non-integrable case. *Comm. Math. Phys.* **120**, 269–294.
- Montgomery, R. [1990] Isoholonomic problems and some applications. *Comm. Math Phys.* **128**, 565–592.
- Montgomery, R. [1991a] *Optimal Control of Deformable Bodies and Its Relation to Gauge Theory* in The Geometry of Hamiltonian Systems, T. Ratiu ed., Springer-Verlag.
- Montgomery, R. [1991b] How much does a rigid body rotate? A Berry’s phase from the 18th century. *Am. J. Phys.* **59**, 394–398.
- Montgomery, R., J.E. Marsden, and T.S. Ratiu [1984] Gauged Lie-Poisson structures, *Cont. Math. AMS* **28**, 101–114.
- Morrison, P.J. [1987] Variational principle and stability of nonmonotone Vlasov-Poisson equilibria. *Z. Naturforsch.* **42a**, 1115–1123.
- 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.
- Murray, R.M. and S. S. Sastry [1993] Nonholonomic Motion Planning: Steering Using Sinusoids. *IEEE Trans. on Automatic Control*, **38**, 700–716.
- Neishtadt, A. [1984] The separation of motions in systems with rapidly rotating phase, *P.M.M. USSR* **48**, 133–139.
- Newcomb, W.A. [1958] Appendix in Bernstein, B. [1958] Waves in a plasma in a magnetic field. *Phys. Rev.* **109**, 10–21.
- Newcomb, W.A. [1962] Lagrangian and Hamiltonian methods in Magnetohydrodynamics. *Nuc. Fusion Suppl.*, part 2, 451–463.
- Oh, Y.G. [1987] A stability criterion for Hamiltonian systems with symmetry, *J. Geom. Phys.* **4**, 163–182.
- Oh, Y.G., N. Sreenath, P.S. Krishnaprasad and J.E. Marsden [1989] The dynamics of coupled planar rigid bodies Part 2: bifurcations, periodic solutions, and chaos, *Dynamics and Diff. Eq’ns.* **1**, 269–298.
- Olver, P.J. [1986] *Applications of Lie groups to differential equations*. Graduate Texts in Mathematics **107**, Springer-Verlag, Berlin.
- O’Reilly, O, N.K. Malhotra and N.S. Namamchchivaya [1996] Some aspects of destabilization in reversible dynamical systems with application to follower forces, *Nonlinear Dynamics*, **10**, 63–87.
- Ortega, J.P. and T.S. Ratiu [1997a] Symmetry, Reduction, and Stability in Hamiltonian Systems *preprint*.
- Ortega, J.P. and T.S. Ratiu [1997b] Stability of relative equilibria. Symplectic block diagonalization. *preprint*.

- Ortega, J.P. and T.S. Ratiu [1997c] Persistence et différentiabilité de l'ensemble des éléments critiques relatifs dans les systèmes hamiltoniens symétriques *C.R. Acad. Sci.* (to appear).
- Ortiz, M. (1986). A note on energy conservation and stability of nonlinear time-stepping algorithms. *Computers and Structures*, **24**, 167–168.
- Ostrowski, J. [1996] *Geometric Perspectives on the Mechanics and Control of Undulatory Locomotion*, PhD dissertation, California Institute of Technology.
- Ostrowski and Burdick [1996] Gait kinematics for a serpentine robot, *IEEE Int. Conf. on Robotics and Automation*, 1294-9, Minneapolis, April, 1996.
- Ostrowski, J., J. W. Burdick, A. D. Lewis & R. M. Murray [1995] The mechanics of undulatory locomotion: The mixed kinematic and dynamic case. *IEEE Int. Conf. on Robotics and Automation*, 1945-1951, Nagoya, Japan, May, 1995.
- Ostrowski, J., J. P. Desai, and V. Kumar. Optimal gait selection for nonholonomic locomotion systems. Submitted to the *IEEE Conf. on Robotics and Automation*, September 1996.
- Ovsienko, V.Y. and B.A. Khesin [1987] Korteweg-de Vries superequations as an Euler equation. *Funct. Anal. and Appl.* **21**, 329–331.
- Patrick, G. [1989] The dynamics of two coupled rigid bodies in three space, *Cont. Math. AMS* **97**, 315–336.
- Patrick, G. [1992] Relative equilibria in Hamiltonian systems: The dynamic interpretation of nonlinear stability on a reduced phase space, *J. Geom. and Phys.* **9**, 111–119.
- Patrick, G. [1995] Relative equilibria of Hamiltonian systems with symmetry: linearization, smoothness and drift, *J. Nonlinear Sci.* **5**, 373–418.
- Pedlosky, J. [1987] *Geophysical Fluid Dynamics*, 2nd Edition, Springer, New York.
- Pedroni, M. [1995] Equivalence of the Drinfeld-Sokolov reduction to a bi-Hamiltonian reduction. *Lett. Math. Phys.* **35**, 291–302.
- Pekarsky, S. and J.E. Marsden [1998] Point Vortices on a Sphere: Stability of Relative Equilibria. *preprint*.
- Poincaré, H. [1885] Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation, *Acta. Math.* **7**, 259.
- Poincaré, H. [1890] *Théorie des tourbillons*, Reprinted by Éditions Jacques Gabay, Paris.
- Poincaré, H. [1890] Sur la problème des trois corps et les équations de la dynamique, *Acta Math.* **13**, 1–271.
- Poincaré, H. [1892–1899], *Les Méthodes Nouvelles de la Mécanique Celeste*. 3 volumes. English translation *New Methods of Celestial Mechanics*. History of Modern Physics and Astronomy **13**, Amer. Inst. Phys., 1993.
- Poincaré, H. [1892] Les formes d'équilibre d'une masse fluide en rotation, *Revue Générale des Sciences* **3**, 809–815.
- Poincaré, H. [1901a] Sur la stabilité de l'équilibre des figures piriformes affectées par une masse fluide en rotation, *Philosophical Transactions A* **198**, 333–373.

- Poincaré, H. [1901b] Sur une forme nouvelle des équations de la mécanique, *CR Acad. Sci.* **132**, 369–371.
- Poincaré, H. [1910] Sur la precession des corps deformables. *Bull Astron* **27**, 321–356.
- Pullin, D.I., and P.G. Saffman [1991] Long time symplectic integration: the example of four-vortex motion, *Proc. Roy. Soc. Lon. A* **432**, 481-494.
- Ratiu. T.S. [1980], *Thesis*. University of California at Berkeley.
- Ratiu. T.S. [1981] Euler-Poisson equations on Lie algebras and the N -dimensional heavy rigid body. *Proc. Natl. Acad. Sci. USA* **78**, 1327–1328.
- Ratiu, T.S. [1982] Euler-Poisson equations on Lie algebras and the N -dimensional heavy rigid body, *Am. J. Math.* **104**, 409–448, 1337.
- Reich, S. [1993]. Symplectic integration of constrained Hamiltonian systems by Runge-Kutta methods. Technical Report 93-13, University of British Columbia.
- Reich, S. [1994] *Symplectic integrators for systems of rigid bodies*. (preprint, Inst. Angw. Anal. Stch.)
- Reich, S. [1994] Momentum preserving symplectic integrators. *Physica D*, 76(4):375–383.
- Reinhall, P.G., T.K. Caughey, and D.Q. Storti [1989] Order and chaos in a discrete Duffing oscillator, implications for numerical integration, *Trans. ASME*, **56**, 162–176.
- Ruth, R. [1983] A canonical integration techniques, *IEEE Trans. Nucl. Sci.* **30**, 2669–2671.
- Sanz-Serna, J. M. [1988] Runge-Kutta shemes for Hamiltonian systems, *BIT* **28**, 877–883.
- Sanz-Serna, J. M. [1991] Symplectic integrators for Hamiltonian problems: an overview. *Acta Num.*, **1**, 243–286.
- Sanz-Serna, J. M. [1996] Backward error analysis of symplectic integrators. *Fields Inst. Comm.*, **10**, 193–205.
- Sanz-Serna, J. M. and M. Calvo [1994] *Numerical Hamiltonian Problems*. Chapman and Hall, London.
- Satzer, W.J. [1977] Canonical reduction of mechanical systems invariant under abelian group actions with an application to celestial mechanics, *Ind. Univ. Math. J.* **26**, 951–976.
- Scovel, C. [1991] Symplectic numerical integration of Hamiltonian systems, *Geometry of Hamiltonian systems*. ed. by T. Ratiu, Springer Verlag, 463–496.
- Simo, J.C. and O. Gonzalez. [1993], Assessment of Energy-Momentum and Symplectic Schemes for Stiff Dynamical Systems. *Proc. ASME Winter Annual Meeting, New Orleans*, Dec. 1993.
- Simo, J.C., D.R. Lewis, and J.E. Marsden [1991] Stability of relative equilibria I: The reduced energy momentum method, *Arch. Rat. Mech. Anal.* **115**, 15-59.
- 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.

- Simo, J.C., J.E. Marsden, and P.S. Krishnaprasad [1988] The Hamiltonian structure of nonlinear elasticity: The material, spatial, and convective representations of solids, rods, and plates, *Arch. Rat. Mech. Anal.* **104**, 125–183.
- Simo, J.C., T.A. Posbergh and J.E. Marsden [1989] Stability analysis of a rigid body with attached geometrically nonlinear rod by the energy-momentum method *Cont. Math. AMS*, **97**, 371-398
- Simo, J.C., T.A. Posbergh, and J.E. Marsden [1990] Stability of coupled rigid body and geometrically exact rods: block diagonalization and the energy-momentum method, *Physics Reports* **193**, 280–360.
- Simo, J.C., T.A. Posbergh, and J.E. Marsden [1991] Stability of relative equilibria II: Three dimensional elasticity, *Arch. Rat. Mech. Anal.* **115**, 61–100.
- Simo, J.C., M.S. Rifai, and D.D. Fox [1992], On a stress resultant geometrically exact shell models. Part VI: Conserving algorithms for nonlinear dynamics. *Comp. Meth. Appl. Mech. Engng* **34**, 117–164.
- Simo, J.C. and N. Tarnow [1992] The discrete energy momentum method. Conserving algorithms for nonlinear elastodynamics, *ZAMP* **43**, 757–792.
- Simo, J.C. and N. Tarnow [1994] A New Energy Momentum Method for the Dynamics of Nonlinear Shells, *Int. J. Num. Meth. Eng.*, **37**, 2527–2549.
- Simo, J.C. and N. Tarnow and M. Doblare [1993] Nonlinear Dynamics of Three-Dimensional Rods: Exact Energy and Momentum Conserving Algorithms. *Comp. Meth. Appl. Mech. Eng.* (to appear).
- Simo, J.C., N. Tarnow, and K.K. Wong [1992] Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics *Comp. Meth. Appl. Mech. Eng.* **100** 63–116.
- Simo, J.C. and L. VuQuoc [1985] Three-dimensional finite strain rod model. Part II. Computational Aspects. *Meth. in Appl. Mech. Eng.*, 58, 79-116.
- Simo, J.C. and L. VuQuoc [1988a] On the dynamics in space of rods undergoing large overall motions—a geometrically exact approach. *Computer Methods in Applied Mechanics and Engineering* **66**, 125-161.
- Simo, J.C. and L. VuQuoc [1988b] The role of nonlinear theories in the dynamics of fast rotating flexible structures. *Journal of Sound and Vibration* **119**, 487-508.
- Simo, J.C. and K.K. Wong [1989] Unconditionally stable algorithms for the orthogonal group that exactly preserve energy and momentum, *Int. J. Num. Meth. Eng.* **31**, 19–52, addendum, **33**, 1321-1323, 1992.
- Sjamaar, R. and E. Lerman [1991] Stratified symplectic spaces and reduction, *Ann. of Math.* **134**, 375–422.
- Smale, S. [1970] Topology and Mechanics, *Inv. Math.* **10**, 305–331, **11**, 45–64.
- Sreenath, N., Y.G. Oh, P.S. Krishnaprasad, and J.E. Marsden [1988] The dynamics of coupled planar rigid bodies. Part 1: Reduction, equilibria and stability, *Dyn. and Stab. of Systems* **3**, 25–49.

- Sternberg, S. [1977] Minimal coupling and the symplectic mechanics of a classical particle in the presence of a Yang–Mills field, *Proc. Nat. Acad. Sci.* **74**, 5253–5254.
- Stofer, D.M. [1987] *Some Geometric and Numerical Methods for Perturbed Integrable Systems*. Thesis, Zurich.
- Sudarshan, E.C.G. and N. Mukunda [1974] *Classical Mechanics: A Modern Perspective*. Wiley, New York, 1974; Second Edition, Krieger, Melbourne–Florida, 1983.
- Tilbury, D., R. Murray, and S. Sastry [1993] Trajectory generation for the N -trailer problem using Goursat normal form. *Proc. IEEE Control and Decision Conf, San Antonio, Dec. 1993*.
- Touma, J. and J. Wisdom [1994] Lie-Poisson integrators for rigid body dynamics in the solar system. *Astro. J.* **107**, 1189.
- Touma, J. and J. Wisdom [1994] Evolution of the Earth-Moon system. *Astro. J.* **108**, 1943.
- Vanhaecke, P. [1996] *Integrable Systems in the Realm of Algebraic Geometry*, Springer Lecture Notes in Mathematics, **1638**.
- Verlet, L. [1967]. Computer experiments on classical fluids. *Phys. Rev.* **159**, 98–103.
- Vershik, A.M. and V. Ya Gershkovich [1988] Non-holonomic Riemannian manifolds, in Dynamical Systems 7, Mathematical Encyclopaedia series 16 (in Russian), MIR pub., translation to be pub. by Springer.
- Veselov, A.P. [1988] Integrable discrete-time systems and difference operators. *Funct. An. and Appl.* **22**, 83–94.
- Veselov, A.P. [1991] Integrable Lagrangian correspondences and the factorization of matrix polynomials. *Funct. An. and Appl.* **25**, 112–123.
- Vierkandt, A. [1892] Über gleitende und rollende Bewegung. *Monatshefte der Math. und Phys.* **III**, 31–54.
- Vinogradov, A.M. and B.A. Kupershmidt [1977] The structures of Hamiltonian mechanics. *Russ. Math. Surv.* **32**, 177–243.
- Wald, R.M. [1993] Variational principles, local symmetries and black hole entropy. *Proc. Lanczos Centenary volume* SIAM, 231–237.
- Walsh, G. and L. Bushnell [1993] Stabilization of multiple input chained form control systems. *Proc. IEEE Control and Decision Conf, San Antonio, Dec. 1993*.
- Walsh, G. and S. Sastry [1995] On reorienting linked rigid bodies using internal motions. *IEEE Trans. on Robotics and Automation*, **11(1)**, 139–146.
- Wang, L.S. and P.S. Krishnaprasad [1992] Gyroscopic control and stabilization, *J. Nonlinear Sci.* **2**, 367–415.
- Weinstein, A. [1978] A universal phase space for particles in Yang-Mills fields, *Lett. Math. Phys.* **2**, 417–420.
- Weinstein, A. [1990] Connections of Berry and Hannay type for moving Lagrangian submanifolds. *Adv. in Math.* **82**, 133–159.

- Wendlandt, J.M. and J.E. Marsden [1997] Mechanical integrators derived from a discrete variational principle, *Physica D* **106**, 223–246.
- Whitham, G.B. [1974] *Linear and Nonlinear Waves*, Wiley-Interscience, pp. 461-462.
- Wisdom, J., and M. Holman [1992] Symplectic maps for the N body problem, *Astron. J.*, **102**, 1528–1538.
- Wisdom, J., and M. Holman and J. Touma [1996] *Symplectic correctors*. *Fields Inst. Comm.*, **10**, 217–244.
- Wisdom, J., S.J. Peale, and F. Mignard [1984] The chaotic rotation of Hyperion, *Icarus* **58**, 137–152.
- Zenkov, D.V. [1995] The Geometry of the Routh Problem, *J. Nonlinear Sci.* **5**, 503–519.
- Zenkov, D.V., A.M. Bloch, and J.E. Marsden [1997] The Energy Momentum Method for the Stability of Nonholonomic Systems *Dyn. Stab. of Systems.*, Dec, 1997.