NOTES FOR LECTURE 3: REDUCTION

HENRY O. JACOBS

ABSTRACT. In this lecture I will discuss Lagrangian reduction. We will first try to formalize the notion of symmetry and discuss what it means for a function, a tensor, or any object to be symmetric with respect to a group of transformations. We will then see what results when the structures of mechanics (symplectic structures, variational principles) hold symmetries. This will allow us to describe dynamics on quotient spaces, the process of which is called "reduction".

1. Lagrangian Mechanics

In this section we will review Lagrangian mechanics and go over Noether's theorem.

2. EL-equations

Theorem 2.1. *The Euler-Lagrange equations:*

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

are equivalent to

$$\delta \int_0^T L(q, \dot{q}) dt = 0$$

for variations with fixed end-points.

Proof. Let $\delta q(t) = \frac{d}{d\epsilon}\big|_{\epsilon=0} q_\epsilon(t)$ be a variation of $q(t)$ induced by the deformation $q_\epsilon(t)$ where $q_o(t) = q(t)$. Fixed endpoints means $\delta q(0) = \delta q(T) = 0$. We find

$$\delta \int_0^T L(q, \dot{q}) dt = \int_0^T \frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} dt$$

using integration by parts , and noting the boundary terms are 0 yields

$$= \int_0^T \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right) \delta q dt$$

If the above is = 0 for all possible variations, δq, then the EL equations must follow. This argument is reversible. ☐

3. Noether's Theorem

Theorem 3.1. *Let G be a Lie group acting on Q and $L : TQ \to \mathbb{R}$ be a Lagrangian such that $L((q,\dot{q}) \cdot g) = L(q,\dot{q})$. Then the quantity*

$$J_\eta := \frac{\partial L}{\partial \dot{q}} \cdot \eta_Q$$

is conserved by the EL equations for each $\eta \in \mathfrak{g} := T_e G$ where η_Q is the vector field on Q given by $\eta_Q(q) = \frac{d}{dt}\big|_{t=0} \exp(t\eta) \cdot q$.

Proof. Take the variation

$$\delta q = \frac{d}{d\epsilon} g_\epsilon \cdot q := \eta \cdot q(t)$$

for some $g_\epsilon \in G$ such that $\eta = \frac{dg_\epsilon}{d\epsilon}$. Then

$$\delta \int_0^T L(q,\dot{q}) dt = 0$$

since $L(q,\dot{q}) = L(g_\epsilon \cdot (q,\dot{q}))$. However

$$\begin{aligned}\delta \int_0^T L(q,\dot{q}) dt &= \int_0^T \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} dt \\ &= \int_0^T \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)\right) \delta q dt + \left(\frac{\partial L}{\partial \dot{q}} \cdot \eta_Q \right)\Bigg|_{t=0}^{t=T}\end{aligned}$$

by the EL-equations this means

$$\frac{\partial L}{\partial \dot{q}} \cdot \eta_Q \Bigg|_{t=0} = \frac{\partial L}{\partial \dot{q}} \cdot \eta_Q \Bigg|_{t=T}$$

□

4. Some Examples

Consider the following examples of physical systems, each of which has a symmetry.

system	**symmetry**	**dynamics**
rigid body in a vacuum	rotational	$\dot{\Pi} = \Pi \wedge \Omega$
particle in a magnetic field	guage-symmetry	$\ddot{x} = \dot{x} \wedge B + \nabla \Phi$
fluid	particle relabeling	$\dot{u} + u \cdot \nabla u = -\nabla p, \mathrm{div}(u) = 0$

This table summarizes the following facts. The first column regards the rigid body in a vacuum (i.e. no gravity). Then if you rotate your head the equations should be the same. Therefore, it should be possible to write the equations in such a way that is independent of how you orient your head. We see this in the right column, where the rigid body equations only use the angular momenta and velocity, but no reference to an initial orientation. Same story for the other two examples. In the next section we will discuss how to understand what symmetry means in a more formal sense.

5. Symmetry and Lie Groups

What is symmetry? We say an equilateral triangle is "symmetric" because we may rotate it 60 degrees and it "looks" the same. This is roughly what we mean by symmetry. Something is symmetric with respect to a transformation, if it "looks" the same after applying the transformation. More generally we will be concerned with groups of transformations. In this case we would say that the equilateral triangle has D_3 symmetry, since we may reflect it about its center and rotate 60 degrees without changing the way it looks. To get beyond this example, much of the work will be in defining what I mean by "looks". To do this we define Lie groups formally and then go through a sequence of increasingly complex examples.

5.1. **Lie Groups.** Let us jump right into the meat and define Lie groups (informally)

Definition 5.1 (informal)**.** *A group (acting on a set M) is a set of transformations from M to M, denoted G, such that:*

1. *G includes the identity transformation e, defined by $e(m) = m$.*
2. *$g_1, g_2 \in G$ implies that $g_1 \circ g_2 \in G$.*
3. *Each element of G is invertible.*

We call G a Lie group *if G is also a smooth manifold, the composition, $\circ$, is a smooth map from $G \times G \to G$, and inversion is smooth from $G \to G$.*

By definining Lie groups we may now state what symmetry is more formally. We could say that an object (defined on M) is symmetric with respect to G if it is unchanged by elements of G. Let's see what this means in some examples.

Example 5.1. *Let $f : M \to \mathbb{R}$ be a smooth real valued function and let G be a Lie group acting on M. Then we say that f is symmetric with respect to G if $f(g \cdot m) = f(m)$ for all $g \in G$.*

Now recall that if M is a manifold, then it has a tangent bundle ,TM, with projection, $\tau_M : TM \to M$. Similarly, we have the dual bundle which is call the cotangent bundle and denoted, T^*M, equipped with the projection $\pi_M : T^*M \to M$. Recall that a one-form is a map $\alpha : M \to T^*M$ such that $\pi_M \circ \alpha(m) = m$ for all $m \in M$. In otherwords, α is a covector field.

Example 5.2. *Let $\alpha \in \Omega^1(M)$ be a one form on M. We can think of α as a real valued function on TM. We say that α is symmetric with respect to G if $\alpha \circ Tg = \alpha$. Given the wedge product of two one forms gives us the set of two forms. We say $\alpha \wedge \beta$ is symmetric with respect to G if α and β are also, taking elements of this form allows us to define G symmetry for 2-forms. Iterating this we can get a definition of G-symmetry for n-forms.*

6. Quotient Spaces

In this section we have the following setup. Let G be a Lie group and M be a manifold. This will allow us to define the quotient space. To do that we first define an equivalence relation. We say that $m_1 \sim_G m_2$ for $m_1, m_2 \in M$ if $m_1 = gm_2$ for some $g \in G$. Given an $m \in M$ the set of all equivalent elements is $G \cdot m = \{gm : g \in G\}$. We call $G \cdot m$ the G-orbit of m. Another popular notation is $[m]_G := G \cdot m$. We can define the set of equivalence classes

$$[M]_G := \{[m]_G \quad | \quad m \in M\}$$

which we call the *quotient of M by G*. Additionally we may define the qoutient projection $\pi : m \in M \to [m]_G \in [M]_G$. We would like the quotient space to be a smooth manifold with π a smooth submersion. However we require some additionally assumptions for this to be the case. We say that G acts *freely* if $g \cdot m = m$ implies $g = e$ for any $m \in M$. Additionally, G is said to act properly if for any convergent sequence $m_k \in M$ then convergence of $g_k m_k \in M$ implies convergence of g_k for any sequence $g_k \in G$. The following proposition roughly says that under suitable conditions the quotient space is a smooth manifold and the quotient projection is a smooth map such that the inverse image is a "copy" of G.

Proposition 6.1. *If G acts freely and properly on M then $\pi : M \to [M]_G$ is a surjective submersion. And $[M]_G$ is a smooth manifold.*

We will not have time to prove this. All quotients in these lectures will be assumed to be of this form though.

Proposition 6.2. *Let G act on M and let $f \in C^\infty(M)$ be a smooth function on M which is G invariant. Then there is a well defined function $[f] \in C^\infty([M])$ such that $[f] \circ \pi = f$.*

Proof. Normally if one applies f to an arbitrary subset of M one should expect to get a subset of $\mathbb{R}$ in the range. However, we will find that f maps each G orbit to an isolated element of $\mathbb{R}$. Recall the definition of G invariance. This means $f(gm) = f(m)$. Noting that this is the case for all $g \in G$ we find that if we apply f to the entire orbit $[m]_G$ yields the single real number $f(m)$. Thus we may define $[f]$ by evaluating f on orbits. □

Similar constructions work for differential forms. If α is an 1-form which is G-invariant than we may define a one-form on M/G in the same fashion (that is by viewing α as a real-valued function).

Question 6.1. Construct a feasible definition of G symmetry for n-forms.

Of particular interest to us is the case where we have a symplectic manifold (P, Ω) where Ω is G invariant under some action on P. A map which preserves Ω is called a symplectomorphism, and one says that G *acts by symplectomorphisms* or simply G acts symplectically.

Lastly, we will not use the following fact explicitly, but it seems natural to ask what it means for a vector field to be G-symmetric. The answer is simply that a vector field X is G symmetric if $Tg \cdot X = X \circ g$.

6.1. **Some facts an notation on Lie groups.** Given a Lie group, G, we may consider it's tangent bundle, TG, in order to talk about velocities on G. In particular we can consider small deviations from the identity. This is the vector space $\mathfrak{g} = T_e G$. We call this vector space the Lie algebra of G. The Lie algebra is associated with a bracket $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ given by the following idea. Let $\xi, \eta \in \mathfrak{g}$. Then we can define a left invariant vector field on G by the assignment $g \mapsto g \cdot \xi$. Similarly for η. Then $[\xi, \eta]$ is the element of $\mathfrak{g}$ which epresses the change in the vector field created by η while flowing along the vector field created by ξ. We see that the Lie bracket satisfies the following two properties

(1) It is anti-symmetric, $[\xi, \eta] = -[\eta, \xi]$.
(2) It satisfies the Jacobi identity. $[[\xi, \eta], \chi] - [\xi, [\eta, \chi]] + [\eta, [\xi, \chi]] = 0$.

Finally, we introduce the ad notation. For each $\xi \in \mathfrak{g}$ we define the map $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ by $\mathrm{ad}_\xi(\eta) = [\xi, \eta]$. Additionally, the dual map to this is $\mathrm{ad}_\xi^* : \mathfrak{g}^* \to \mathfrak{g}^*$. Specifically this mean, given a covector $\alpha \in \mathfrak{g}^*$ and a vector $\eta \in \mathfrak{g}$ then ad_ξ^* is defined by the condition $\langle \alpha, \mathrm{ad}_\xi(\eta) \rangle = \langle \mathrm{ad}_\xi^*(\alpha), \eta \rangle$.

7. Lagrangian Reduction

By Noether's theorem and the above discussion two things should be obvious now. Given a G invariant Lagrangian, $L : TQ \to \mathbb{R}$ there should exists a well defined object on $[TQ]_G$. Secondly, Noether's theorem suggests that the equations of motion secretly evolve on a lower dimensional space. In fact we will find that this space is $[TQ]_G$. To do this in it's full glory for general spaces requires more than a lecture (see the first three sections of [2]). However, we are able to handle the most important example, which is that of a G invariant Lagrangian defined on TG. In this case $TG/G = \mathfrak{g}$ and we get equations of motion on the Lie algebra. These eqautions are called the Euler-Poincaré equations and this type of reduction is called Euler-Poincaré reduction.

8. Euler-Poincaré

Reduction on Lie groups. Let $L : TG \to \mathbb{R}$ and say that it is G symmetric. What then happens?

Theorem 8.1. *Let $L : TG \to \mathbb{R}$ be a left G invariant Lagrangian on a Lie group G. Let $g(t)$ be a curve in G and $\delta g(t)$ a variation of a curve in G. Then the following are equivalent.*

(1) *g satisfies EL equations.*
(2) *g extremizes the action S for an arbitrary variaion δg with fixed end points.*

(3) $\xi(t) = g^{-1}\dot{g}$ *extremizes the reduced action* $[S]$ *with respect to variation* $\delta\xi = \dot{\eta} + [\eta, \xi]$ *for an arbitrary curve* $\eta(t) \in \mathfrak{g}$ *which is* 0 *at the end points.*
(4) $\xi(t) = g^{-1}\dot{g}$ *satisfies the Euler-Poincaré equation*

$$\frac{d}{dt}\left(\frac{\partial[L]}{\partial\xi}\right) = \mathrm{ad}_{\xi}^{*}\left(\frac{\partial[L]}{\partial\xi}\right)$$

Proof. Equivalence of the first two is obvious. For the equivalence of (2) and (3) we can start by calculating the variation in ξ induced by the variaiton in g. Let $g(t, \epsilon)$ be a deformation of $g(t)$ and set $deltag = \frac{\partial g}{\partial \epsilon}$ be the variation and define $\eta = g^{-1}\delta g$. We see firstly that:

$$\begin{aligned} \delta\xi &= \delta(g^{-1}\dot{g}) \\ &= -g^{-1} \cdot \delta g \cdot g^{-1} + g^{-1}\frac{\partial^2 g}{\partial g \partial\epsilon} \\ &= -\eta\xi + g^{-1}\frac{\partial^2 g}{\partial t\partial\epsilon} \end{aligned}$$

Similarly we calculate

$$\begin{aligned} \dot{\eta} &= \frac{d}{dt}(g^{-1}\delta g) \\ &= -g^{-1}\dot{g}g^{-1}\delta g + g^{-1}\frac{\partial^2 g}{\partial t\partial\epsilon} \\ &= -\xi\eta + \frac{\partial^2 g}{\partial t\partial\epsilon} \end{aligned}$$

As a result we see that $\delta\xi - \dot{\eta} = \xi\eta - \eta\xi$. In otherwords $\delta\xi = \dot{\eta} + [\xi, \eta]$. If this is the case we see that

$$\begin{aligned} \delta\int L(g, \dot{g})dt &= \delta\int L(g^{-1}(g, \dot{g}))dt \\ &= \delta\int [L](\xi(t))dt \end{aligned}$$

However, the last integral is extremized with respect to variations of ξ of the desired form. Finally we establish the equivalence of (3) and (4). Assume (3). Then:

$$\delta\int [L](\xi)dt = \int \langle\frac{\partial[L]}{\partial\xi}, \dot{\eta} + [\xi, \eta]\rangle dt$$

using integration by parts on the time derivative of η we find

$$= \int \langle -\frac{d}{dt}\left(\frac{\partial[L]}{\partial\xi}\right), \eta\rangle + \langle \frac{\partial[L]}{\partial\xi}, \mathrm{ad}_\xi(\eta)\rangle dt$$
$$= \int \langle -\frac{d}{dt}\left(\frac{\partial[L]}{\partial\xi}\right) + \mathrm{ad}^*_\xi\left(\frac{\partial[L]}{\partial\xi}\right), \eta\rangle dt$$

Since η is arbitrary, the integrand must be 0. This implies the Euler-Poincaré equations. Again, this reasoning goes both ways, so that we've proven equivalence. □

The above construction was for left actions. To understand the difference between left and right actions requires a more thourough investigation of Lie groups than we have time for [5, see §9]. However, for those who are aware of the difference, right action yields the variation is $\delta\xi = \dot\eta + [\eta, \xi]$ and the EP equations is $\frac{d}{dt}\frac{\partial[L]}{\partial\xi} = -\,\mathrm{ad}^*_\xi\left(\frac{\partial[L]}{\partial\xi}\right)$.

8.1. **Some examples.** The rigid body is given given by a Lagrangian $L : T\,\mathrm{SO}(3) \to \mathbb{R}$,

$$L(R, \dot R) = \mathrm{trace}\left((R^{-1}\dot R)^T \mathbb{I} R^{-1}\dot R\right)$$

EP reduction yields the equations of motion

$$\dot\Pi = \Pi\Omega - \Omega\Pi$$

where $\Pi = \mathbb{I}\cdot\Omega$ and $\Omega = R^{-1}\dot R$. By identifying the set of antisymmetric 3x3 matrices with $\mathbb{R}^3$ this becomes the normal rigid body equation that we are used to. In particular, it would be written in terms of the cross product on $\mathbb{R}^3$. This is an example that one should master to go further and I recommend the first part of [**?**] to see this done slowly and carefully.

The second example of importance is the Lie group $\mathrm{SDiff}(M)$ consisting of volume preserving diffeomorphisms of a Riemannian manifold M. Given by

$$L(\varphi, \dot\varphi) = \frac{\rho}{2}\int_M \|\dot\varphi(x)\|^2 d\mathrm{vol}.$$

We see that this L has right invariant symmetry with respect to $\mathrm{SDiff}(M)$ and so the dynamics satisfy the EP equations. The Lie algebra of $\mathrm{SDiff}(M)$ is the set of divergence free vector fields, and so the EP dynamics are given by an evolution equation on the set of divergence free vector fields. In the case that M is a subset of $\mathbb{R}^d$ we write these as

$$\frac{\partial u}{\partial t} + u\cdot\nabla u = -\frac{1}{\rho}\nabla p$$
$$\mathrm{div}(u) = 0.$$

Those trained in fluid dynamics should recognize these equations as the inviscid fluid equations. To get the Navier-Stokes equations one must add a viscous force.

The intrinsic formulation of the viscous term is known, but (unfortunately) is usually not addressed in most GM literature. To see this in great detail we refer to the final chapter of [5]. To study specifically the case of the rigid body and fluids in one book see [1].

9. Symplectic Reduction

In Symplectic Hamiltonian mechanics we have two main objects. Symplectic forms and Hamiltonians. We just illustrated in the last section how G-symmetric objects induce equivalent objects on the quotient space. The process of finding these objects in the context of symplectic Hamiltonian mechanics is called Symplectic reduction (or reduction by symmetry). To do this takes a little more work than one might initially think though. We first must understand what a momentum map is.

Definition 9.1. *Let (P, Ω) be a symplectic manifold and let Ω be G invariant. A map $\mathbf{J} : P \to \mathfrak{g}^*$ such that $i_{\xi_P}\Omega = d\langle \mathbf{J}, \xi\rangle$ for all $\xi \in \mathfrak{g}$ is called a* momentum map. *We say that $\mathbf{J}$ is* equivariant *if $\mathbf{J} \circ g = \mathrm{Ad}^*_g \circ \mathbf{J}$ for all $g \in G$.*

Finding momentum maps is usually a simple matter of writing down the definition and massaging the equations into a form one prefers. I will attempt to find a good example to compute in the lecture. In any case given a symplectic manifold, (P, Ω), with a symplectic G-action, a momentum map, $\mathbf{J}$, and a G-invariant Hamiltonian, H, we call the tuple $(P, \Omega, G, \mathbf{J}, H)$ a Hamiltonian G-space.

Proposition 9.1. *Given a Hamiltonian G space, $(P, \Omega, G, \mathbf{J}, H)$, the momentum map is conserved under the evolution of Hamilton's equations.*

Proof. This is the Hamiltonian version of Noether's theorem. One simply needs to calculate the time derivative of $\mathbb{J}$ in the natural way. Let $\xi \in \mathfrak{g}$ be a constant element of the Lie algebra. Then

$$\frac{d}{dt}\langle \mathbf{J}, \xi\rangle = X_H[\langle \mathbf{J}, \xi\rangle] = d\langle \mathbf{J}, \xi\rangle \cdot X_H = \Omega(\xi_P, X_H)$$

However, by the definition of X_H this last quantity is $\Omega(\xi_P, X_H) = -\langle dH, \xi_P\rangle$. We can write the last term as $\frac{d}{d\lambda}(H \circ \exp(\lambda\xi))$. But since H is G-invariant this last term is 0. By allowing ξ to be arbitrary we see that $\frac{d}{dt}\mathbf{J} = 0$ □

This makes it clear that the dynamics of a Hamiltonian G space evolve on a level set $\mathbf{J}^{-1}(\mu)$ for some $\mu \in \mathfrak{g}^*$. However, this is not the end of the story. It turns out that just because Ω is a symplectic form on P does not imply that it is a symplectic form on the subset $\mathbf{J}^{-1}(\mu)$. This is illustrated most clearly in the symplectic reduction theorem (which we will not prove)

Theorem 9.1. *Let (P, Ω) be a symplectic manifold and G be a Lie group which acts on on P by symplectomorphisms. Given a equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$*

there exists a symplectic form on the space $P_\mu := \mathbf{J}^{-1}(\mu)/G_\mu$ *where* $G_\mu := \{g \in G : g \cdot \mu = \mu\}$. *This symplecic form,* Ω_μ *on* P_μ *is defined by the property that*

$$\pi_\mu^* \Omega_\mu = i_{mu}^* \Omega_P$$

Where $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ *and* $i_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow P$.

For more information see [4] and sources therein (particularly the first two chapters).

10. A final note for tomorrow's Lecture

How do we simulate these systems (fluids and rigid bodies are things we want to simulate). Here is an idea. Consider the action functional $S : Q \times Q \to \mathbb{R}$ defined by

$$S(q_1, q_2; h) = \text{ext}\left(\int_0^h L(q, \dot{q})dt\right)$$

Given a curve $q(t)$ on the interval $[0, nh]$ which satisfies the EL equations, a variation (without fixed end points) of S yields nothing except the boundary terms.

$$\begin{aligned} 0 &= \delta S(q_0, q_n) \\ &= \sum_{k=0}^{n-1} \delta S(q_k, q_{k+1})[q] \\ &= \sum_{k=0}^{n-1} D_1 S(q_k, q_{k+1})[q]\delta q_k + D_2 S(q_k, q_{k+1})[q]\delta q_{k+1} \\ &= \sum_{k=0}^{n-1} (D_1 S(q_k, q_{k+1})[q] + D_2 S(q_{k-1}, q_k)[q])\delta q_k \end{aligned}$$

for arbitrary variations δq with fixed endpoints $q(0), q(nh)$ and $q_k = q_{k+1}$ and $\delta q_k = \delta q(kh)$. We call the equation

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

the *discrete Euler-Lagrange equation*. We see that $S(q_0, q_n)$ satisfies the discrete Euler-Lagrange equations along exact solution of the continuous Euler Lagrange equations for L. If one view's the above formula as a method for obtainingin q_{k+1} from q_k and q_{k-1} then we could consider this as a integrator for the equations of motion. More on this will be discussed tomorrow. For more information see [6] and see [3] for the Hamiltonian description.

References

[1] V I Arnold and B A Khesin. *Topological Methods in Hydrodynamics*, volume 24 of *Applied Mathematical Sciences*. Springer Verlag, 1992.

[2] H Cendra, J E Marsden, and T S Ratiu. *Lagrangian Reduction by Stages*, volume 152 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 2001.

[3] E Hairer, C Lubich, and G Wanner. *Geometric Numerical Integration, Structure Preserving Algorithms for Ordinary Differential Equations*, volume 31 of *Series in Computational Mathematics*. Springer Verlag, 2002.

[4] J E Marsden, G Misiolek, J P Ortega, M Perlmutter, and T S Ratiu. *Hamiltonian Reduction by Stages*, volume 1913 of *Lecture Notes in Mathematics*. Springer, 2007.

[5] J E Marsden and T S Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer Verlag, 2nd edition, 1999.

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