Simple mechanical Systems with Symmetry

Lars Sydnes

Master of Science in Physics and Mathematics
Submission date: June 2007
Supervisor: Eldar Straume, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

Following Arnold [1], in classical mechanics there is the important class of simple mechanical systems on Riemannian manifolds whose symmetry group G is a group acting by isometries on the configuration manifold M. Many of the basic ideas in the study of such systems date back to Euler, Lagrange, Jacobi, Routh, Riemann and Poincaré. However, nowadays these ideas can also be developed in a more modern differential geometric setting.

In general, the symmetry group is used to reduce the equations of motion to a simpler system of differential equations, for example at the level of the orbit space M/G.

[1] Arnold, V.I. Mathematical Methods of Classical Mechanics. Second edition, Graduate Texts in Mathematics 60, Springer-Verlag, 1989.

Assignment given: 13. February 2007
Supervisor: Eldar Straume, MATH

Simple mechanical systems with symmetry

Lars Sydnes

ABSTRACT. We go through the basic theory of simple mehcanical systems with symmetry, in an attempt to understand some of the main features of configuration space reduction. As a part of this, we will look at some special cases for whom this works out well, and also indicate a direction of further development.

Preface

This project started when I studied a treatment of the classical three body problem [**6**] given by Hsiang and Straume. I understood that they used Riemannian geometry. However, to gain computational clarity they hided this behind a lot of vector analysis. As a simple method of book keeping, I tried to formulate as much as I could of it explicitly in terms of smooth manifolds and Riemannian geometry. This task lead me away from that three body problem, and into the land of what we will call Simple mechanical systems.

There exists some books touching this subject (Oliva [**9**], Bloch [**2**] and Bullo-Lewis [**5**]), but I do not know books treating this subject in its own right. In my opinion, it is a pity, since Riemannian geometry fits very well together with classical mechanics. In the beginning of the work with this project I tried to avoid the literature, discovering the theory on my own. Later in the process I confered the literature to control my own results and to learn more.

There is a huge literature on classical mechanics in general, focusing on the Lagrangian and Hamiltonian approach. In my work I have tried to follow another approach, where I separate out the pure geometrical features. In my opinion, this sheds light over the relationship between classical mechanics and geometry. I aimed at a description purely in terms of Riemannian geometry, hiding every variational principle behind the Riemannain connection. However, I must admit that as the project evolved, I got into serious difficulties in carrying this out, and I had to use some variational arguments to work around those problems. So I learned that the standard approach is very powerful in some situations, and to respect the main-stream approaches to classical mechanics, a field that has occupied thousands of bright heads for more than 300 years.

In the beginning of the work with this project I tried to avoid the literature, discovering the theory on my own. Later in the process I confered the literature to control my own results and to learn more.

When it comes to the reduction theory, I was very optimistic half a year ago. I stated and proved a very powerful theorem that seemingly did not exist in the literature. Ofcourse, the result was not was not true, but I believed in it until I found a small error in the proof about a month ago(and I could not believe that I ever believed in that theorem). The ghosts of this theorem still live in this text. However, it is broken down into special cases.

I want to thank my family for the support during my studies the last five years in Trondheim.

I will also thank my advisor Eldar Straume for his kind guidance for the last three years, allowing me to go in both the right and wrong directions.

Contents

1. Introduction

1.1. Simple mechanical systems. The mechanical principles that Newton formulated in his *Principia* depended heavily the assumption that the space we live in can be described in terms of Euclidean geometry. Later, it was developed formalisms, like the Lagrange and Hamilton- formalisms, that were independent of the choice of coordinates. I intend to demonstrate that Riemannian geometry is a convenient setting for a coordinate-free formulation of many problems in classical mechanics.

However, when Riemannian geometry is introduced, the equations of motion suddenly reappears as a Riemannian counterpart of Newton's second law

$$\text{Force} = \text{mass} \times \text{acceleration}.$$

The reason for this is that the Riemannian geometry can be used to define covariant acceleration as well as a notion of mass, where the usual scalar mass is replaced by a tensorial quantity, the kinematic metric.

EXAMPLE 1.1. Consider a system consisting of two point masses m_1, m_2 moving in plane. If we choose a coordinate system, the configuration of such a system may be described by four real numbers (x_1, y_1, x_2, y_2), where (x_i, y_i) is the position of particle nr. i. We assume that the forces between the particles is modeled by a potential function $U(x_1, y_1, x_2, y_2)$ (Equal to −potential energy), such that we can write the equations of motion as

$$\begin{aligned} m_i \ddot{x}_i &= \frac{\partial U}{\partial x^i} \\ m_i \ddot{y}_i &= \frac{\partial U}{\partial y^i} \end{aligned} \tag{2}$$

This can be rewritten in terms of Riemannian geometry. Consider the Riemannian manifold $(M, \mathfrak{m}) = (\mathbb{R}^4, \mathfrak{m})$, where

$$\mathfrak{m} = m_1(dx_1^2 + dy_1^2) + m_2(dx_2^2 + dy_2^2),$$

measuring the kinetic energy of motions. For a curve $\gamma(t) = (x_1(t), y_1(t), x_2(t), y_2(t))$, the covariant acceleration of such a curve, $\ddot{\gamma} = \nabla_{\dot{\gamma}}\dot{\gamma} = (\ddot{x}_1, \ddot{x}_1, \ddot{x}_2, \ddot{y}_2)$ since the Christoffels symbols vanish because the coefficients of the Riemannian metric are constant.

The gradient of U with respect to this Riemannian geometry will be the vector

$$\text{grad } U = \left(\frac{1}{m_1}\frac{\partial U}{\partial x_1}, \frac{1}{m_1}\frac{\partial U}{\partial y_1}, \frac{1}{m_2}\frac{\partial U}{\partial x_2}, \frac{1}{m_2}\frac{\partial U}{\partial y_2} \right).$$

Hence, we can write equation (2) as

$$\ddot{\gamma} = \text{grad } U$$

Fortunately, this equation is valid for any smooth coordinate system on M. However, it is hard to believe that we can write the equations of motion in a more convenient way than equation (2).

The advantage of Riemannian geometry will be clear first at the point where we want to impose constraints on the system. Suppose for example that we have a reason to believe that $x_1^2 + y_1^2 + x_2^2 + y_2^2 = L$ (For example if the point masses are joined by a wire of fixed length, passing through the origin.). Then, the motion will be restricted to the 3-sphere $\mathbb{S}^3 \subseteq \mathbb{R}^4$. Later we will see how we restrict the original mechanical system to such submanifold. Anyway, we get a system consisting of a Riemannian metric on $\mathbb{S}^3$ and the restriction $\tilde{U} = U|_{\mathbb{S}^3}$. The geometry on the sphere will prevent us from writing the equations of motion explicitly in an elegant way like equation (2). However, the equation

$$\nabla_{\dot{\gamma}}\dot{\gamma} = \text{grad } \tilde{U}$$

is still valid.

This example is easily generalized to more complex systems. And such examples has given me the personal opinion that Riemannian geometry is a very convenient setting for understanding of classical mechanics. However, if I want to write down the equations of motion in a coordinate system (q_i), I will write down the Lagrangian function

$$\mathcal{L} = \frac{1}{2}\mathfrak{m}(\dot{q}_i, \dot{q}_i) + \tilde{U}(q_i),$$

and then solve the Euler-Lagrange equations

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

belonging to the Lagrangian formalism.

We will treat this topic in the second half of the first chapter. The first half of that chapter is devoted to the differential geometric background.

1.2. Symmetry. In the real life, a mechanical problem hardly appears without any symmetry present. In our case such symmetries will be represented by smooth Lie group actions on smooth manifolds. From the theory of dynamical systems, we know that such symmetries can be used to reduce the dimension of the system. Ofcourse, this applies directly to classical mechanical systems. But, we would rather do this in a more geometric way, allowing the geometry to illuminate the process.

We will discover Noether's theorem, the theorem tying together continous symmetries and conservation laws. This theorem has a very simple proof in our setting. However, the main task will be to understand the following problem: Given a mechanical system with configuration manifold M and a Lie group G of smooth symmetries: Formulate the equations of motion in terms of a mechanical system on the orbit space M/G (which may or may not be a manifold) and a method of lifting of curves in M/G to M.

Symmetries and conservation laws will be treated in the end of the first chapter, while the reduction problem occurs throughout the second chapter.

CHAPTER 1

Simple mechanical systems

1. Riemannian geometry

Here we review some well known facts from differential geometry. I rely mainly on the presentations given by Sipvak in [**12**] and do Carmo [**4**]. The part explaining the action of the Lie derivative on the tensor algebra follows partially Kobayashi-Nomizu [**7**], with some small differences concering the grading of the tensor algebra.

1.1. Smooth manifolds. In this thesis we will work in the category of finite dimensional smooth manifolds and smooth maps. Hence, if an object or morphism can be specified to be smooth, the reader can safely assume that we mean that it is smooth.

Manifolds will typically be denoted by the letters M, N. The $\mathbb{R}$-algebra of smooth real valued functions on a manifold M is denoted by $\mathcal{F}(M)$, or simply $\mathcal{F}$, when the choice of M is obvious.

The tangent and cotangent bundle of the manifold M will be denoted respectively by TM and T^*M. Some other vector bundles will be denoted by $\mathcal{H}M$, $\mathcal{V}M$. For a given (smooth, by default) vector bundle $\mathcal{V}M \to M$, the $\mathcal{F}$-module of (smooth) sections will be denoted by $\mathfrak{X}[\mathcal{V}M]$. The fibre of $\mathcal{V}M$ over $x \in M$ will be denoted by $\mathcal{V}_xM$. Hence, T_xM will denote the tangent space of M at $x \in M$.

For a given vector bundle $\mathcal{V}M \to M$, the $\mathcal{F}$-module of sections will be denoted by $\mathfrak{X}[\mathcal{V}M]$. However, we use the shorthands $\mathfrak{X}(M) = \mathfrak{X}[TM]$ and $\mathfrak{X}^*(M) = \mathfrak{X}[T^*M]$.

Given a smooth map $f : M \to N$, there is an associated smooth bundle map $Tf : TM \to TN$, the derivative of f. This derivative follows the chain rule: $T(f \circ g) = Tf \circ Tg$. This map is the union of fibre maps $T_xf : T_xM \to T_{f(x)}N$.

Given a map $f : M \to N$ there is an $\mathcal{F}$-module $\mathfrak{X}^f(M)$, consisting of vector fields X on M such that there is a vector field Y on M with $Y_{f(x)} = Tf(X_x)$. In particular, we note that if f is a diffeomorphism, $\mathfrak{X}^f(M) = \mathfrak{X}(M)$. We use the notation $Tf : \mathfrak{X}^f(M) \to \mathfrak{X}(N)$ for the assignment $X \mapsto Y$.

We have a pairing $\mathfrak{X} \otimes_{\mathbb{R}} \mathcal{F} \to \mathcal{F}$, given by $\langle X, f\rangle(x) = X_x(f)$, the derivative of f along the tangent vector X_x at $x \in M$. This pairing is natural in the sense that when we have a map

$$M \xrightarrow{f} N,$$

inducing

$$\mathcal{F}(N) \xrightarrow{f^*} \mathcal{F}(M) \qquad g \mapsto g \circ f$$

$$\mathfrak{X}^f(M) \xrightarrow{Tf} \mathfrak{X}(N) \qquad X \mapsto TfX$$

in such a way that

$$\langle X, g^*f\rangle = \langle X, f \circ g\rangle = g^*\langle TgX, f\rangle$$

whenever $X \in \mathfrak{X}^g(M)$ and $g \in \mathcal{F}(N)$. This is clear from the definition of the push forward of tangent vectors

We also have a natural pairing $\langle -, -\rangle : \mathfrak{X} \otimes \mathfrak{X}^* \to \mathcal{F}$ that allows us to identify $\mathfrak{X}^*$ with the $\mathcal{F}$-dual of $\mathfrak{X}$. Any map $f : M \to N$ will induce a map $f^* : \mathfrak{X}^*(N) \to \mathfrak{X}^*(M)$. f^* is defined by

$$\langle X, f^*\omega\rangle(x) = \langle (T_x f)X_x, \omega\rangle,$$

and hence this pairing is also natural in the above sense.

The vector space $\mathfrak{X}(M)$ has a natural Lie algebra structure coming from the commutator

$$[-, -] : \mathfrak{X}(M) \otimes_{\mathbb{R}} \mathfrak{X}(M) \to \mathfrak{X}(M), \quad \langle [X, Y], f\rangle = \langle X, \langle Y, f\rangle\rangle - \langle Y, \langle X, f\rangle\rangle.$$

We call this operation the *Lie-bracket*. This operation posses a naturality similar to the naturality of the pairing of vector fields and functions, i.e.

$$Tg[X, Y] = [TgX, TgY] \quad \text{whenever} \quad X, Y \in \mathfrak{X}^g(M).$$

This follows directly from the corresponding property of the paring of vector fields and functions.

Along with the $\mathcal{F}$-module $\mathfrak{X}$, we have the $\mathcal{F}$-dual $\mathfrak{X}^*$, and also the $\mathcal{F}$-tensor-algebra $\tau(M)$ of tensor fields on M. I define this is a G-graded $\mathcal{F}$-algebra, where G is the semigroup of words generated by two letters, say u, d, where the group operation is concatenation of words. However, it is usual to group covariant and contravariant components, and thus construct an algebra graded over $\mathbb{N}^2$, like in Kobayashi-Nomizu [**7**]

We define this algebra recursively: For two words w_1, w_2, we define the module in w_1w_2-grade in $\tau(M)$, $\tau^{w_1w_2}(M)$ to be equal to $\tau^{w_1}(M) \otimes_{\mathcal{F}} \tau^{w_2}(M)$ up to natural isomorphism, and in particular

$$\tau^{\emptyset}(M) = \mathcal{F}(M) \quad \tau^d(M) = \mathfrak{X}(M) \quad \tau^u(M) = \mathfrak{X}^*(M),$$

where $\emptyset$ denotes the empty word. We define the multiplication in $\tau(M)$ by the identity function

$$\tau^{w_1}(M) \otimes_{\mathcal{F}} \tau^{w_2}(M) \to \tau^{w_1w_2}(M)$$

Given a word w with a u in the i-th position and a d in the j-th position we can write $w = w_1uw_2dw_3$ or $w = w_1dw_2uw_3$, and we get a word $c^i_jw = w_1w_2w_3$, simply by deletion of the indicated letters. Associated with this, there is a map

$$C^i_j : \tau^w(M) \to \tau^{c^i_j w}(M)$$

defined by

$$C^i_j(\omega_1 \otimes \sigma_1 \otimes \omega_2 \otimes \sigma_2 \otimes \omega_3) = \langle \sigma_1, \sigma_2\rangle \omega_1 \otimes \omega_2 \otimes \omega_3$$

This map is called the i, j-contraction. The symbol C^i_j may or may not be applicable in the different degrees.

Now, we are in position to define the Lie derivative $\mathfrak{L}_X$ associated with the vector field X. This is the unique derivation in the $\mathcal{F}$-algebra $\tau(M)$ commuting with contractions and preserving degrees, ie,

$$\mathfrak{L}_X(\omega \otimes \sigma) = (\mathfrak{L}_X\omega) \otimes \sigma + \omega \otimes (\mathfrak{L}_X\sigma)$$

$$\mathfrak{L}_X(C^i_j\omega) = C^i_j(\mathfrak{L}_X\omega)$$

such that

$$\mathfrak{L}_X f = \langle X, f \rangle \quad \text{and} \quad \mathfrak{L}_X Y = [X, Y].$$

The Lie derivative may also be defined as the derivative of tensor fields along the flow of the vector field. Ie, if $\omega \in \tau^{u^n}(M)$ and X is an infinitesimal generator of the local flow θ_t, then

$$\mathfrak{L}_X \omega = \frac{d}{dt} \theta_t^* \omega \mathcal{V}_{t=0}$$

We say that ω is *constant* along X when $\mathfrak{L}_X \omega = 0$.

1.2. Connections. An affine connection ∇ on a vector bundle $\mathcal{V}M$ over a manifold M is a $\mathbb{R}$-linear map

$$\nabla : \mathfrak{X}(M) \otimes_{\mathbb{R}} \mathfrak{X}[\mathcal{V}M] \to \mathfrak{X}[\mathcal{V}M],$$

such that $\nabla_X fY = X(f) + f\nabla_X Y$ and $\nabla_{fX} Y = f\nabla_X Y$, whenever $X \in \mathfrak{X}, f \in \mathcal{F}, Y \in \mathfrak{X}[\mathcal{V}M]$.

An affine connection ∇ in the tangent bundle TM is called *symmetric* if $\nabla_X Y - \nabla_Y X = [X, Y]$.

1.2.1. *Pullback of connections.* Suppose that $f : N \to M$ is a map, and that $\mathcal{V}M \to M$ is a smooth vector bundle equipped with a linear connection ∇. By the usual pullback-construction, we get a smooth vector bundle $f^*(\mathcal{V}M) \to N$.

Now, if $Y \in \mathfrak{X}[f^*\mathcal{V}M]$ and $X \in \mathfrak{X}(N)$, we want to define a section $\nabla_X Y \in \mathfrak{X}[f^*\mathcal{V}M]$: For a given point $p \in N$, there is a neighbourhood U such that $f : U \to M$ is an embedding. By choosing U small enough, we ensure that the restrictions $X|U, Y|U$ are extendable to smooth vector fields $\tilde{X}, \tilde{Y}$ on M. Then we define $\nabla_X Y(p) = \nabla_{\tilde{X}} \tilde{Y}(f(p))$. This is independent of choice of extensions $\tilde{X}, \tilde{Y}$, which can be verified by inspection of the formula for $\nabla_{\tilde{X}} \tilde{Y}$ in a well chosen coordinate system. Thanks to the local character of smoothness, this gives us a smooth section $\nabla_X Y \in \mathfrak{X}[f^*\mathcal{V}M]$.

The linearity and derivation-property of this operation is preserved during this pullback. Hence, we get a linear connection

$$\nabla : \mathfrak{X}(N) \otimes_{\mathbb{R}} \mathfrak{X}[f^*\mathcal{V}M] \to \mathfrak{X}[f^*\mathcal{V}M]$$

If $\mathcal{V}M$ is the tangent bundle TM, every $X, Y \in \mathfrak{X}(N)$ can be regarded as sections in $f^*(TM)$. Hence, we get a map

$$\nabla : \mathfrak{X}(N) \otimes_{\mathbb{R}} \mathfrak{X}(N) \to \mathfrak{X}[f^*TM].$$

In the case that N is an interval $I \subseteq \mathbb{R}$, this construction gives us the covariant derivative along curves: For a regular curve $\gamma : I \to M$, a section of γ^*TM is essentially the same as a function $V : I \to TM$ such that $V(t) \in T_{\gamma(t)}M$. On I we have the special vector field ∂_t, coming from the oriented Riemannian structure on I. In this situation, there is a section $\nabla_{\partial_t} V \in \mathfrak{X}[\nabla]$. We refer to this as the *covariant derivative* of V along γ, and use the shorthands $\nabla_{\dot{\gamma}} V$ or simply $\dot{V}$ for this vector field. In the continuation of this, we establish the notation $\ddot{\gamma}$ as a replacement of $\nabla_{\dot{\gamma}} \dot{\gamma}$, and for a covector field $p \in \mathfrak{X}^*[\gamma^*TM]$, we write $\dot{p}$ instead of $\nabla_{\partial_t} p$.

1.3. Riemannian metrics. Informally, *Riemannian metric* $\mathfrak{m}$ on a manifold M is a smooth family of inner products $\mathfrak{m}_x$ on the tangent spaces T_xM. More formal, a Riemannian metric $\mathfrak{m}$ is a 2-tensor $\mathfrak{m} \in \tau^{uu}(M) = \mathfrak{X}^*(M) \otimes \mathfrak{X}^*(M)$ such that $\mathfrak{m}(X,Y) = \mathfrak{m}(Y,X)$ and $\mathfrak{m}(X,X)(x) > 0$ whenever $X_x \neq 0$. Possesing such a device, we can measure lengths of a tangent vectors as well as angles between them. Integration of lengths of velocity vectors gives a notion of length along curves.

We can regard a Riemannian metric $\mathfrak{m}$ as a positive definite, symmetric linear map

$$\mathfrak{m} : \mathfrak{X}(M) \otimes_{\mathcal{F}} \mathfrak{X}(M) \to \mathcal{F}(M).$$

This map characterises $\mathfrak{m}$ completely.

In the same way as an inner product on a vector space V gives a canonical isomorphism $V \to V^*$, a Riemannian metric gives a vector bundle isomorphism

$$\tilde{\mathfrak{m}} : TM \to T^*M, \quad \tilde{\mathfrak{m}}(v_x)(w_x) = \mathfrak{m}(v_x, v_y),$$

called the *inertia operator*. We can obviously recover the metric $\mathfrak{m}$ from the associated inertia operator $\tilde{\mathfrak{m}}$.

1.3.1. *The Riemannian gradient.* On any manifold there is a natural derivation $d : \mathcal{F} \to \mathfrak{X}^*$, $f \mapsto df$, defined by $\langle X, f\rangle = \langle X, df\rangle$.

On a Riemannian manifold, the inertia operator $\tilde{\mathfrak{m}}^{-1} : \mathfrak{X}^* \to \mathfrak{X}$ gives a canonical derivation $\blacktriangledown = \tilde{\mathfrak{m}}^{-1} \circ d : \mathcal{F} \to \mathfrak{X}$. The vector field $\blacktriangledown f$ is called the *gradient vector field* of the function f. The gradient is characterized by

$$\mathfrak{m}(X, \blacktriangledown f) = \langle X, df\rangle = \langle X, f\rangle.$$

1.4. The Riemannian connection. On a Riemannian manifold $(M, \mathfrak{m})$ there is a unique symmetric affine connection ∇ on the tangent bundle TM such that

$$X(\mathfrak{m}(Y,Z)) = \mathfrak{m}(\nabla_X Y, Z) + \mathfrak{m}(Y, \nabla_X Z)$$

for $X, Y, Z \in \mathfrak{X}(M)$. We say that ∇ is compatible with $\mathfrak{m}$.

∇_X may be extended to a $\mathbb{R}$-linear endomorphism the tensor algebra $\tau(M)$. We indicate only the first step. Any covector field ω on M can be represented by a vector field Y such that $\omega = \tilde{\mathfrak{m}}Y$. Then we may define

$$\nabla_X \omega = \nabla_X(\tilde{\mathfrak{m}}Y) = \tilde{\mathfrak{m}}(\nabla_X Y),$$

even though this is usually stated as a consequence of the definition of $\nabla_X : \tau(M) \to \tau(M)$ that emphasises that ∇_X is a derivation commuting with contractions.

2. Riemannian submersions

In this section we will investigate some properties of submersions of manifolds. Starting with some general features depending only on the smooth structure, we will gradually introduce more and more geometry.

2.1. Surjective submersions. By the term *submersion*, we will mean *surjective submersion*, i.e. a surjective map $\pi : M \to N$ (of smooth manifolds) such that $T_x\pi : T_xM \to T_{\pi(x)}N$ is surjective for every $x \in M$.

We define the *vertical subspace* at $x \in M$, $\mathcal{V}_xM = \ker(T_x\pi)$. This gives us a family $\mathcal{V}M \subseteq TM$ of subspaces of the tangent spaces of M. $\mathcal{V}M$ is a smooth distribution with constant rank. This is easily seen since $\mathcal{V}M$ corresponds to the smooth foliation $\{\pi^{-1}y\}_{y\in N}$ of M .

At $x \in M$, we get an isomorphism $T_xM/\mathcal{V}_xM \cong T_{\pi(x)}N$. Then, if we choose a smooth sub-bundle $\mathcal{H}M \subseteq TM$ complementary to $\mathcal{V}M$, ie, a bundle such that

$$TM \cong \mathcal{V}M \oplus \mathcal{H}M,$$

then $TM/\mathcal{V}M \cong \mathcal{H}M$, and hence, $T_x\pi : \mathcal{H}_xM \to T_{\pi(x)}N$ is an isomorphism. We call $\mathcal{H}M$ a chosen *horizontal* distribution.

Along with the decomposition $T = \mathcal{V} \oplus \mathcal{H}$, the projections $\mathrm{pr}_{\mathcal{H}} : T \to \mathcal{H}$ and $\mathrm{pr}_{\mathcal{V}} : T \to \mathcal{V}$ can be regarded as smooth tensors of type $\binom{1}{1}$. Each of them determine the decomposition, since $\ker(\mathrm{pr}_{\mathcal{V}}) = \mathcal{H}M$, $\ker(\mathrm{pr}_{\mathcal{H}}) = \mathcal{V}M$ and $\mathrm{pr}_{\mathcal{V}} = 1 - \mathrm{pr}_{\mathcal{H}}$

2.1.1. *Horizontal liftings.* Consider a submersion $M \to N$ and a vector field $X \in \mathfrak{X}(N)$. A vector field $Y \in \mathfrak{X}(M)$ is a *lifting* of X if and only if

$$T_x\pi(Y_x) = X_{\pi(x)} \tag{1}$$

for all $x \in M$.

Obviously, such liftings are not in general unique: Assume that $Y_1, Y_2 \in \mathfrak{X}(M)$ are lifitngs of $X \in \mathfrak{X}(N)$. As we see from equation (1), this implies that

$$T_x\pi((Y_1 - Y_2)_x) = 0,$$

which is the same as $(Y_1 - Y_2)_x \in \mathcal{V}M$. Conversely, you can add a vertical vector field $Y_1 \in \mathfrak{X}(\mathcal{V}M)$ to a lifting $Y \in \mathfrak{X}(M)$ of a vector field $X \in \mathfrak{X}(N)$, and the result $Y + Y_1$ will also be a lifting of X. Hence, lifting of vector fields is a map

$$\mathrm{L} : \mathfrak{X}(N) \to \mathfrak{X}[TM/\mathcal{V}M]$$

If we have defined a horizontal distribution $\mathcal{H}M$, such that $TM = \mathcal{H}M \oplus \mathcal{V}M$, we get a canonical isomorphism $\mathrm{C} : TM/\mathcal{V}M \cong \mathcal{H}M$, and hence, we get a map

$$\mathrm{L}_0 : \ \mathfrak{X}(N) \xrightarrow{\mathrm{L}} \mathfrak{X}[TM/\mathcal{V}M] \xrightarrow{\mathrm{C}} \mathfrak{X}[\mathcal{V}M]$$

This defines *horizontal lifting of vector fields*. This can be described in terms of lifting of single tangent vectors $v \in TN$:

Let $v \in T_{\pi(x)}N$ be a vector tangent to N. Let $\mathrm{L}_{0x}v$ be the inverse image of v through the isomorphism $T_x\pi : \mathcal{H}_xM \cong T_{\pi(x)}N$. We call $\mathrm{L}_{0x}v$ the *horizontal lifting* of v at $x \in M$. For a vector field $X \in \mathfrak{X}(N)$, we note that L_0X on M is given by $(\mathrm{L}_0X)_x = \mathrm{L}_{0x}X_{\pi(x)}$.

Now, we have almost proved the following result:

LEMMA 2.2. *If X is a smooth vector field on N, then the horizontal lifting* L_0X *is the unique horizontal smooth vector field on M that is a lifting of X.*

PROOF. We let $X \in \mathfrak{X}(N)$. I must prove that L_0X is a smooth vector field on M . The rest is clear from the discussion above. Obviously L_0X is a section of the tangent bundle $TM \to M$, hence we only need to prove that L_0X is smooth at any point $x \in M$

Let $x \in M$. From the properties of submersions, we know that there exist a neighbourhood U of x such that $\pi : U \to \pi(U)$ is diffeomorphic to the projection

$$\pi : \mathbb{R}^n \oplus \mathbb{R}^{m-n} \to \mathbb{R}^n$$

on the first factor. On this neighbourhood, we get a coordinate-dependent rule $\tilde{H}$ for lifting of vector fields, defined by

$$\tilde{H}(\frac{\partial}{\partial x^i}) = \frac{\partial}{\partial x^i},$$

such that for any smooth vector field $X \in \mathfrak{X}(U)$ we get a smooth vector field $\tilde{H}X \in \mathfrak{X}(U)$. This vector field has the property that $T_x\pi((\tilde{H}X)_x) = X_{\pi(x)}$.

Furthermore, the projection $pr_{\mathcal{H}}$ onto $\mathcal{H}M$ is a smooth tensor field on M of type $\binom{1}{1}$. Hence, the restriction of $pr_{\mathcal{H}}$ to U will applied to $\tilde{H}X$ give a smooth vector field $pr_{\mathcal{H}}(\tilde{H}X)$ contained in the distribution $\mathcal{H}U = \mathcal{H}M|U$. $pr_{\mathcal{H}}(\tilde{H}X) - \tilde{H}X \in \mathcal{V}U$, hence is $\pi_{*x}((pr_{\mathcal{H}}(\tilde{H}X))) = T_x\pi((\tilde{H}X)_x) = X_{\pi(x)}$. Thus, $pr_{\mathcal{H}}(\tilde{H}X) = \mathrm{L}_0X|_U$, and hence, L_0X is smooth on the neighbourhood U of $x \in M$. □

As noted above, general liftings of vector fields differ form horizontal ones by vertical vector fields. Now, if we are given a fixed vertical vector field $V \in \mathfrak{X}(\mathcal{V}M)$, we can define the V-lifting, $L_V\mathfrak{X}(N) \to \mathfrak{X}(M)$ by the formula $L_V(X) = \mathrm{L}_0X + V$. We call $L_V(X)$ the V-lifting of X.

2.1.2. *Lifting of functions and tensors.* It is simple to lift a map $f : N \to P$ to M. We define $\mathrm{L}_0f = \pi^*f = f \circ \pi$. Similarly, the pullback of ccovectorfields give an obvious map $\mathrm{L}_0 : \mathfrak{X}^*(N) \to \mathfrak{X}^*(M)$, $\mathrm{L}_0\omega = \pi^*\omega$. We have the following properties

LEMMA 2.3. *For $X \in \mathfrak{X}(N)$, $\omega\mathfrak{X}^*(N)$ and $f \in \mathcal{F}(N)$ we have*

$$\langle \mathrm{L}_0X, \mathrm{L}_0\omega\rangle = \mathrm{L}_0\langle X, \omega\rangle \quad \text{and} \quad \langle \mathrm{L}_0X, \mathrm{L}_0f\rangle = \mathrm{L}_0\langle X, f\rangle$$

PROOF. This is a direct consequence of the naturality of the pairings involved. $T\pi(\mathrm{L}_0X) = X$, and hence is

$$\langle \mathrm{L}_0X, \mathrm{L}_0\omega\rangle = \langle \mathrm{L}_0X, \pi^*\omega\rangle = \pi^*(\langle T\pi(\mathrm{L}_0X), \omega\rangle) = \mathrm{L}_0\langle X, \omega\rangle$$

By substitution $\omega \leftrightarrow f$ we get the second part of the lemma. □

The horizontal lifting extends to a well behaved lifting $\mathrm{L}_0 : \tau(N) \to \tau(M)$, taking tensor fields on N to tensor fields on M.

2.1.3. *The Lie bracket.* The horizontal lifting L_0X of a vector field $X \in \mathfrak{X}(N)$ obviously belongs to $\mathfrak{X}^\pi(M)$, and $T\pi\mathrm{L}_0X = X$. From the naturality of the Lie bracket we get the following important fact, relating the Lie bracket on N to that of M:

LEMMA 2.4.

$$\mathrm{L}_0[X, Y] = \mathrm{pr}_{\mathcal{H}}[\mathrm{L}_0X, \mathrm{L}_0Y] \quad X, Y \in \mathfrak{X}(N),$$

where $pr_{\mathcal{H}}$ is the projection onto the horizontal subspaces

PROOF. For vector fields $X, Y \in \mathfrak{X}(N)$,

$$= [T\pi(\mathrm{L}_0X), T\pi(\mathrm{L}_0X)] = T\pi[\mathrm{L}_0X, \mathrm{L}_0Y] \tag{5}$$

However, $\mathrm{L}_0 \circ T\pi = \mathrm{pr}_{\mathcal{H}}$. Hence

$$\mathrm{L}_0[X, Y] = \mathrm{pr}_{\mathcal{H}}[\mathrm{L}_0X, \mathrm{L}_0Y],$$

and the result is proved. □

2.1.4. *Lifting of curves.* From the lifting of tangent vectors, we proceed to lifting of smooth curves. This is essentially a question of integration of lifted velocity vectors. We look at a non-homogeneous lifting.

LEMMA 2.6 (tittel). *Let δ be a curve in N with domain $D_\delta = [a, b]$, let $V \in \mathfrak{X}[\mathcal{V}M]$ be a vertical vector field on M , and let $x \in \pi^{-1}(\delta(a))$ Then, there exists a unique curve γ_V in M such that*

$$\dot{\gamma}_V(t) = \mathrm{L}_{0\gamma(t)}\dot{\delta}(t) + V_{\gamma(t)} := (L_V)_{\gamma(t)}\dot{\delta}(t)$$

and $\gamma_v(a) = x$

PROOF. This is a direct consequence of the existence and uniqueness theorem for solution curves of first order differential equations. Locally, on an interval $(t_0-\varepsilon, t_0+\varepsilon)$, we can extend $(L_V)_{\gamma(t)}\dot{\delta}(t)$ to a vector field on an open set containing $\gamma(t_0)$, except when $\dot{\delta}(t_0) = 0$, and we get $\gamma|_{(t_0-\varepsilon,t_0+\varepsilon)}$ as an integral curve of this vector field. The case when $\dot{\delta}(t_0) = 0$ is different, but we will not go into the details here. $\square$

2.2. Riemannian submersions. Now, we look at the special case of a submersion $\pi : M \to N$ where M is a Riemannian manifold with metric m. In this case, there is a fairly canonical choice of a horizontal distribution, namely $\mathcal{H}M = \mathcal{V}M^\perp \subseteq TM$. $\mathcal{H}M$ will be a smooth distribution since $\mathcal{V}M$ is smooth and the Gram-Schmidt-procedure is smooth. Hence, we have the horizontal lifting for free, and also the results above.

2.2.1. *The induced Riemannian metric.* A Riemannian metric $\mathfrak{m}$ on a manifold M is completely characterized by the associated map $\mathfrak{m} : \mathfrak{X}(M) \otimes \mathfrak{X}(M) \to \mathcal{F}(M)$, taking the pair X, Y of smooth vector fields on M to the smooth function $\mathfrak{m}(X,Y) \in \mathcal{F}(M)$.

The horizontal lifting $\mathrm{L}_0 : \mathfrak{X}(N) \to \mathfrak{X}(M)$ will induce a map $\bar{\mathfrak{m}}_\pi : \mathfrak{X}(N) \otimes \mathfrak{X}(N) \to \mathcal{F}(M)$ by the composition

$$\mathfrak{X}(N) \otimes \mathfrak{X}(N) \xrightarrow{\mathrm{L}_0 \otimes \mathrm{L}_0} \mathfrak{X}(M) \otimes \mathfrak{X}(M) \xrightarrow{\quad\mathfrak{m}\quad} \mathcal{F}(M)$$

Now, for vector fields $X, Y \in \mathfrak{X}(N)$, it may or may not be the case that $\tilde{\mathfrak{m}}_\pi(X,Y)$ can be regarded as a function on N, but when this is the case, $\tilde{\mathfrak{m}}_\pi(X,Y)$ is contained in the image of the injective map $\mathrm{L}_0 : \mathcal{F}(N) \to \mathcal{F}(M)$, and hence, we get a unique smooth function $\mathfrak{m}_\pi(X,Y)$ on N determined by

$$\mathrm{L}_0(\mathfrak{m}_\pi(X,Y) = \tilde{\mathfrak{m}}_\pi(X,Y).$$

In this case the linear map

$$\mathfrak{m}_\pi : \mathfrak{X}(N) \otimes \mathfrak{X}(N) \to \mathcal{F}(N)$$

will determine a Riemannian metric $\mathfrak{m}_\pi$ on N. $\mathfrak{m}_\pi$ is obviously symmetric and bilinear, and also positive definite since the horizontal lifting $\mathrm{L}_0 : \mathfrak{X}(N) \to \mathfrak{X}(M)$ is injective.

We have a nice condition for determining if this construction is possible: Recall that we have an inner product on every tangent space T_xM on M, orthogonal decompositions $T_xM = \mathcal{H}_xM \oplus \mathcal{V}_xM$ and canonical linear isomorphisms $\mathcal{H}_xM \cong T_{\pi(x)}N$. For two points $y, z \in \pi^{-1}(x)$, we get another canonical linear isomorphism $\phi_{yz} : \mathcal{H}_yM \to \mathcal{H}_zM$, the composition $\mathcal{H}_yM \cong T_{\pi(y)}N = T_{\pi(z)} \cong \mathcal{H}_zM$. For all $w \in M$, $\mathcal{H}_wM$ is equipped with an inner product, by restriction of the Riemannian metric. Hence, it makes sense to ask if ϕ_{yz} is an isometry.

LEMMA 2.7. *Let* $(M, \mathfrak{m})$ *be a Riemannian manifold and* $\pi : M \to N$ *a surjective submersion. There exists a unique Riemannian metric* $\mathfrak{m}_\pi$ *on* N *induced by* π *if and only if* $\phi_{yz} : (\mathcal{H}_yM, \mathfrak{m}_x) \to (\mathcal{H}_zM, \mathfrak{m}_z)$ *is an isometry whenever* $\pi(y) = \pi(z)$.

DEFINITION 2.8. Let $(M, \mathfrak{m})$ be a Riemannian manifold and suppose that $\pi : M \to N$ is submersion. π is called a *Riemannian submersion* if the composition

$$\mathcal{H}_yM \cong T_{\pi(y)}N = T_{\pi(z)} \cong \mathcal{H}_zM$$

is an isometry whenever it is defined. The metric $\mathfrak{m}_\pi$ on N defined by

$$\mathfrak{m}_\pi(X, Y) = \mathfrak{m}(\mathrm{L}_0 X, \mathrm{L}_0 Y) \quad X, Y \in \mathfrak{X}(N)$$

is called the Riemannian metric on M *induced* by π, or simply the *induced metric*.

For a pair $((M, \mathfrak{m}), (N, \mathfrak{m}_N))$ of Riemannian manifolds, we define a submersion $\pi : M \to N$ to be a *Riemannian submersion* of the Riemannian manifolds M, N if $\mathfrak{m}_N = \mathfrak{m}_\pi$.

Remark 2.9. In the case of Riemannian submersions, the horizontal lifting can be defined in another perhaps more natural way. Recall the isomorphisms $\tilde{\mathfrak{m}} : \mathfrak{X}(M) \to \mathfrak{X}^*(M)$, $\tilde{\mathfrak{m}}_\pi : \mathfrak{X}(N) \to \mathfrak{X}^*(N)$ and the pullback map $\pi^* : \mathfrak{X}^*(N) \to \mathfrak{X}^*(M)$. Now, we have the diagram

$$\begin{array}{ccc} \mathfrak{X}(N) & \xrightarrow{\tilde{\mathfrak{m}}_\pi} & \mathfrak{X}^*(N) \\ \downarrow{\scriptstyle \mathrm{L}_0} & & \downarrow{\scriptstyle \pi^*} \\ \mathfrak{X}(M) & \xrightarrow{\tilde{\mathfrak{m}}} & \mathfrak{X}^*(M) \end{array},$$

expressing that L_0 is essentially the same map as π^*.

2.2.2. *The gradient.* We now want to express the gradient coming from the metric $\mathfrak{m}_\pi$ on N. We use the notation $\overset{\pi}{\blacktriangledown}$ for this, and recall that it is defined by $\mathfrak{m}_\pi(\overset{\pi}{\blacktriangledown} f, X) = X(f)$ for $X \in \mathfrak{X}(N)$ and $f \in \mathcal{F}(N)$.

The vector field $\blacktriangledown \mathrm{L}_0(f)$ is horizontal since the function $\mathrm{L}_0(f)$ is constant along the fibres of π. Indeed, if V is a vertical vector field, $V(\mathrm{L}_0(f)) = 0$, and hence $\mathfrak{m}(V, \blacktriangledown \mathrm{L}_0(f)) = 0$. Also $\mathrm{L}_0 \overset{\pi}{\blacktriangledown} f$ is horizontal, hence, we can compare $\blacktriangledown \mathrm{L}_0 f$ and $\mathrm{L}_0 \overset{\pi}{\blacktriangledown}$ by looking at the inner product with horizontal vectors.

If $X \in \mathfrak{X}(N)$, then $\mathfrak{m}(\blacktriangledown \mathrm{L}_0 f, \mathrm{L}_0 X) = \mathrm{L}_0 X(\mathrm{L}_0 f) = \mathrm{L}_0(X(f)) = \mathrm{L}_0(\mathfrak{m}_N(\overset{\pi}{\blacktriangledown} f, X) = \mathfrak{m}(\mathrm{L}_0(\overset{\pi}{\blacktriangledown} f), \mathrm{L}_0 X)$.

This implies that

$$\blacktriangledown \mathrm{L}_0 = \mathrm{L}_0 \overset{\pi}{\blacktriangledown}$$

2.2.3. *The connection.* For a Riemannian submersion $\pi : (M, \mathfrak{m}) \to N$, we have Riemannian connections ∇ and $\overset{\pi}{\nabla}$ associated with the Riemannian metrics $\mathfrak{m}$ and $\mathfrak{m}_\pi$ on M, N.

Lemma 2.10. *For vector fields X, Y on N,*

$$\overset{\pi}{\nabla}_X Y = T\pi(\nabla_{\mathrm{L}_0 X} \mathrm{L}_0 Y),$$

Proof. This follows from a long but simple calculation that depends heavily on the way the horizontal lifting works together with inner products and Lie brackets. □

3. G-spaces

3.1. Smooth actions on manifolds. A snooth G-space is a manifold with a smooth action of a Lie group G, $\varphi : G \to \mathrm{Diff}(M)$, $\quad g \mapsto \varphi_g$ such that the corresponding map

$$\varphi : G \times M \to M \quad \varphi(g, m) = \phi_g(m)$$

is smooth. If the map

$$G \times M \to M \times M \quad (g, m) \mapsto (m, gm)$$

is proper, then we say that G acts *properly* on M.

For a point $m \in M$, we have the subgroup $G_m \subseteq G$, consisting of group elements g with $gm = m$. We call G_m the *isotropy* group at m.

Conjugacy gives an equivalence relation $\sim$ on the set of subgroups of G: $H \sim K$ if there exists a $g \in G$ with $gHg^{-1} = K$. We denote the conjugacy class of the subgroup $H \subseteq G$ by (H).

We define the isotropy type of the point m to be the conjugacy class (G_m) of the isotropy group G_m. The set of points in M with isotropy type (H) will be denoted by $M_{(H)}$. Since $G_{gm} = gG_mg^{-1}$, we know that every $M_{(H)}$ is a G-space. For a proper action, the isotropy subgroups are compact.

For a given point $m \in M$, we have a map $\psi_m : G \to M$ given by $\psi_m(g) = \varphi_g(m)$. The image $Gm = \psi_m(G)$ is called the G-orbit of m. The set of G-orbits constitute a partition of M.

If we assume that G acts properly, we can use the slice theorem to prove there there is a smooth manifold structure on the set $M_{(H)}/G$ of G-orbits in $M_{(H)}$ such that $\pi : M_{(H)} \to M_{(H)}/G$ is a submersion. Also following from the slice theorem is the fact that there exists a unique orbit type (H) such that $M_{(H)}$ is an open and dense submanifold of M. Therefore, when we work with proper actions, it is often convenient to discard the set $M - M_{(H)}$.

Hence, we fix the setting of the following discussion: Lie groups act properly on manifolds, with only one isotropy type. I .e. there is subgroup $H \subseteq G$ such that $M_{(H)} = M$.

3.2. Riemannian G-spaces. Under the restriction we made in the last section, associated with G-space, there is a surjective submersion

$$\pi : M \to N = M/G.$$

Hence, we can define the vertical distribution $\mathcal{V}M$ with respect to the action of G. This distribution consists of precisely the vectors that are tangent to the G-orbits in M.

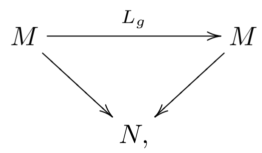
THEOREM 3.1. *When M is a Riemannian manifold with kinematic metric $\mathfrak{m}$ and G acts by isometrics, the submersion $\pi : M \to N$ will be a Riemannian submersion.*

We recall the decomposition $TM \cong \mathcal{H}M \oplus \mathcal{V}M$, where $\mathcal{V}M \perp \mathcal{H}M$, coming from the submersion π and the Riemannian metric $\mathfrak{m}$. This decomposition is clearly invariant under the induced action of G on the tangent bundle TM, and hence, $\mathcal{H}M$ and $\mathcal{V}M$ are G-vector bundles.

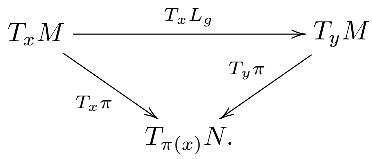
PROOF. Assume that $x, y \in M$ such tat $\pi(x) = \pi(y)$. We must prove that $\phi_{x,y}$ defined by the diagram

(2)
$$\begin{array}{ccc} \mathcal{H}_xM & \xrightarrow{\phi_{xy}} & \mathcal{H}_yM \\ & {\scriptstyle T_x\pi}\searrow^{\cong} \quad {}^{\cong}\swarrow{\scriptstyle T_y\pi} & \\ & T_{\pi(x)}N & \end{array}$$

is an isometry of inner product spaces. Since $\pi(x) = \pi(y)$, there exists a $g \in G$ such that $y = gx = L_g x$, so that the commutative diagram



and behold, there is an induced commuttative diagram



Since $\mathcal{H}M$ is G-invariant, and T_xL_g is defined to be an isometry of inner product spaces, we see that there is an induced map $T_xL_g : \mathcal{H}_xM \to \mathcal{H}_yM$, which is also an isometry. Moreover, this map is equal to ϕ_{xy} since T_xL_g fills diagram (2) in the same way as ϕ_{xy}. □

3.3. The geometry of the orbit space. From the above theorem, we see that we can apply the machinery of Riemannian submersions to our new situation. Hence, we get a Riemannian metric $\mathfrak{m}_N$ on N and a corresponding Riemannian connection $\overset{N}{\nabla}$. In this setting, it is normal to call $\mathfrak{m}_N$ the orbital distance metric. The name is well chosen, since this metric measures locally the minimal distance between G-orbits.

4. Simple mechanical systems

Here follows an exposition of the basic properties of simple mechanical systems. Initially, this is inspired by a treatment of the 3-body problem given by Hsiang-Straume in [**6**]. This can be found in my project thesis "Natural Lagrangian systems on Riemannian manifolds" written in 2006. Another treatment of this topic is Oliva [**9**]. When it comes to the general Lagrangian formalism in classical mechanics, I use Arnold [**1**] as a main reference.

4.1. Lagrangian systems. Now we approach the physics. We will look at a special type of Lagrangian systems on smooth manifolds. A Lagrangian system on a manifold M is determined by a function $\mathcal{L} : TM \to \mathbb{R}$. In this context, $\mathcal{L}$ is called the *Lagrange function* of the Lagrangian system $(M, \mathcal{L})$.

One reason for doing this definition is that we can use $\mathcal{L}$ to characterize a certain class of curves on M, the so called *motions* of the system. They are meant to model physical behaviour, in the sense that the points in the space M represents physical configurations, while a motion γ represents physical change of configuration, ie, physical motions in the space of configurations.

First we need to consider some calculus of variations. Let $\Omega(M)$ denote the set of smooth curves γ in M defined on compact intervals $I_\gamma \subseteq \mathbb{R}$. Given two points $x, y \in M$ and an interval $[a, b] \subseteq \mathbb{R}$, we let $\Omega_x^y(M; [a, b])$ denote the set of smooth curves $\gamma : [a, b] \to M$ such that $\gamma(a) = x$ and $\gamma(b) = y$.

A map $\varphi : (-\epsilon, \varepsilon) \to \Omega(M)$, $s \mapsto \varphi_s$ is called a smooth 1-parameter family of curves if the map $(s, t) \mapsto \varphi_s(t)$ is smooth on its domain of definition.

A smooth 1-parameter family φ of curves, that goes into the subset $\Omega_x^y(M;[a,b]) \subseteq \Omega(M)$ is called a family keeping endpoints fixed. Given a curve $\gamma \in \Omega_x^y(M;[a,b])$, a family $\varphi : (-\epsilon, \varepsilon) \to \Omega_x^y(M;[a,b])$ with $\varphi_0 = \gamma$ is called a variation of γ keeping endpoints fixed.

Along with $\mathcal{L}$, there is a function $\Lambda : \Omega(M) \to \mathbb{R}$ given by

$$\Lambda[\gamma] = \int_{I\gamma} \mathcal{L}(\dot{\gamma})dt$$

For a smooth 1-parameter family φ of curves defined on $s \in (-\epsilon, \varepsilon)$, Λ gives a function $s \mapsto \Lambda[\varphi_s]$.

Λ is said to be *stationary* at γ if $\frac{d}{ds}\Lambda[\varphi_s]\mathcal{V}_{s=0} = 0$ for all variations φ of γ keeping endpoints fixed.

γ is said to be a *motion* of $(M, \mathcal{L})$ if Λ is stationary at γ. It is possible to deduce a set of second order differential equations that completely characterize the motions of the system. They are usually expressed as

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0 \quad i = 1, \dots n, \tag{1}$$

where $(q^1, \dots q^n, \dot{q}^1, \dots, \dot{q}^n)$ is a local coordinate system on TM. This set of equations is called the *Euler-Lagrange*-equations.

4.2. Simple mechanical systems. A simple mechanical system is a Lagrangian system such that there exist a Riemannian metric $\mathfrak{m}$ and a smooth function U on M such that

$$L(v) = \frac{1}{2}\mathfrak{m}(v, v) + U(p)$$

when $v \in T_pM$. In this case we call $\mathfrak{m}$ the *kinematic metric* of the system, and U is called the *potential function of the system.* We also define a quadratic form

$$T = \frac{1}{2}\mathfrak{m}(v, v),$$

called the *kinetic energy.* $V = -U$ is called the *potential energy*

We can heuristically derive the equation that we call *Newton's equation* from the Euler-Lagrange equations.

In a local coordinate system $(q^i, \dot{q}^i)$ on TM, we can write the Lagrangian function on the form

$$\mathcal{L}(q^i, \dot{q}^i) = \frac{1}{2}g_{ij}(q^i) + U(q^i)$$

Using the summation convention, we write the Euler-Lagrange equations as

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

Fixing a point x in the coordinate system, we may, without loss of generality assume that the coordinate system is chosen in such a way that the Christoffels symbols Γ_{ij}^k all vanish at x. In this case,

$$\frac{\partial L}{\partial q^i}dq^i = dU$$

and

$$(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i})dq^i = \nabla_{\dot{\gamma}}\tilde{\mathfrak{m}}\dot{\gamma},$$

and the substitution $p = \tilde{\mathfrak{m}}\dot{\gamma}$ yields (1) on the form

$$\dot{p} = dU. \tag{2}$$

This equation is called *Newton's equation*. The mass operator transforms Newton's equation in to the equivalent forms

$$\ddot{\gamma} = \blacktriangledown U, \quad \tilde{\mathfrak{m}}\ddot{\gamma} = dU, \quad F = \tilde{\mathfrak{m}}\ddot{\gamma},$$

where F is the one-form dU. We will not distinguish between the different forms of this equations. All of them will be called *Newton's equation* whenever it is convenient. The last form of the equation suggests that we can generalize the notion of simple mechanical systems to include systems consisting of a Riemannian manifold $(M, \mathfrak{m})$ together with a 1-form $F \in \mathfrak{X}^*(M)$, representing a generalized force. The motions of such systems will then be curves γ such that

$$F = \tilde{\mathfrak{m}}\ddot{\gamma}$$

Remark 4.3. (Generalized forces). This can be generalized further, to include forces represented by smooth functions $F : TM \to T^*M$.

One option is to introduce dissipative forces, depending on the velocity, modeling air resistance and similar phenomena. A possible approach to this is to model such forces by symmetric bilinear forms $R \in \mathfrak{X}^* \otimes \mathfrak{X}^*$.

I believe that we also can model magnetic forces, represented by alternating 2-forms ω, such that the resulting force $F_\omega : TM \to T^*M$ is given by $F_\omega(v_p)(w_p) = \omega(v_p, w_p)$, and the acceleration $\ddot{\gamma}$ is always perpendicular to $\dot{\gamma}$. This can also be incorporated into the Lagrangian setting. If we let U be a smooth function (representing the electric potential) and A be a 1-form, we can consider the Lagrangian function

$$\mathcal{L}(v_x) = \frac{1}{2}\mathfrak{m}(v_x, v_x) + A(v_x) + U(x), \quad v_x \in T_xM$$

The variational principle will then give the equations

$$\langle \tilde{\mathfrak{m}}\ddot{\gamma}, v\rangle = dA(\dot{\gamma}, v) + dU(v)$$

However, in this thesis we will consider only the case where F can be represented by a 1-form on M, and with some few exceptions F will be the differential dU of a smooth function U. When we emphasize the force F, we will talk about the simple mechanical system $(M, \mathfrak{m}, F)$, whose motions are given by $\dot{p} = F$.

Remark 4.4. (Conservation of energy). For a simple mechanical system with generalized force F, we compute the change in kinetic energy along a fixed motion γ:

$$\frac{dT}{dt} = \frac{d}{dt}\frac{1}{2}\mathfrak{m}(\dot{\gamma}, \dot{\gamma}) = \mathfrak{m}(\dot{\gamma}, \ddot{\gamma}) = F(\dot{\gamma}),$$

Introducing arch length parametrization s, and denoting by $'$ the derivative with respect to arch length, we get $\gamma' = \dot{\gamma}t'$, and along

$$dT = \frac{dT}{dt}dt = \frac{dT}{dt}\frac{dt}{ds}ds = F(\dot{\gamma})t'ds = F(\gamma')ds = \gamma^*F,$$

ie, the 1-form on the interval D_γ measuring the change of kinetic energy is the pullback of F through the curve itself. This implies that the change in kinetic energy along a segment $\gamma[a, b]$ of a motion of the system,

$$T_b - T_a = \int_{[a,b]} \gamma^*F := \int_\gamma F.$$

However, if $F = dU$ for a function U, then this integral is independent of choice of path between $\gamma(a)$ and $\gamma(b)$, and hence we can write

$$T_b - T_a = \int_{\gamma(a)}^{\gamma(b)} dU = U(\gamma(b)) - U(\gamma(a)).$$

This proves that $T_a - U(\gamma(a)) = T_b - U(\gamma(b))$. Consequently, the quantity $T - U$ is conserved along motions of the system.

4.3. Systems with holonomic constraints. Now we will learn how to construct simple Mechanical systems.

In some cases we want to restrict the motions of a system $(M, \mathcal{L})$ to a submanifold $i : N \hookrightarrow M$. If we believe that there are forces keeping the motions within N and that these forces are perpendicular to N, the motions within N will be modeled by a simple mechanical system $(N, i\mathcal{L}')$, where $\mathcal{L}'$ is the restriction of $\mathcal{L}$ to $TN \subseteq TM$. The kinematic metric $\mathfrak{m}_N$ is the pullback $i^*\mathfrak{m}$ and the potential function $\overset{N}{U}$ on N is the restriction i^*U of U to N.

The tangent bundle of M is pulled back to a vector bundle

$$i^*(TM) \downarrow N = TN \oplus \mathfrak{n}N,$$

where $\mathfrak{n}N$ is the orthogonal complement of $TN \subseteq i^*(TM) \downarrow N$ with respect to the inner product imported from TM. $\mathfrak{n}N$ is called the n*normal bundle* of N in M. Together with this, we have the projections pr_T onto TM and $\mathrm{pr}_\mathfrak{n}$ onto the normal bundle.

The gradient $\overset{N}{\blacktriangledown}$ on the Riemannian manifold $(N, \mathfrak{m}_N)$ is given by $\overset{N}{\blacktriangledown} f = \mathrm{pr}_\mathfrak{n} \blacktriangledown \tilde{f}$, where $\tilde{f}$ is an arbitrary extension of f to M and pr.

Recall from section 1.2.1 that there is a connection $i^*\nabla$ on i^*TM, the pullback of the Riemannian connection. The Riemannian connection $\overset{N}{\nabla}$ on N is given by

$$\overset{N}{\nabla}_X Y = \mathrm{pr}_T i^* \nabla_X Y.$$

Hence, the covariant acceleration $\overset{N}{\nabla}_{\dot\gamma} \dot\gamma$ of a curve in N is simply the projection of the acceleration $\nabla_{\dot\gamma}\dot\gamma$ measured in M onto the tangent bundle of N.

The gradient $\overset{N}{\blacktriangledown}$ on the Riemannian manifold $(N, \mathfrak{m}_N)$ is given by $\overset{N}{\blacktriangledown} f = \mathrm{pr}_\mathfrak{n} \blacktriangledown \tilde{f}$, where $\tilde{f}$ is an arbitrary extension of f to M and pr.

Hence, Newton's equations on $(N, \mathcal{L}')$ is simply the projection of Newton's equations in the system $(M, \mathcal{L})$ onto TN. The normal part, that is projected away, will give information about the forces that is needed to keep the motion within N, ie, the constraint forces.

Now we can use this to see how simple mechanical systems arise from a lot of physical situations. A lot of mechanical systems can be modeled as a collection of, say N point particles with different masses $\{m_i \mathcal{V} i = 1, \ldots N\}$ moving in three dimensional Euclidean space. Hence, we get a configuration space $M = \mathbb{R}^{3N}$. Sometimes we may model the interaction between the particles by a potential function U on M. However, different degrees of rigidity may or may not be of interest. Rigidity may be modeled by "approximately infinitely strong" springs joining the different particles. Such a point of view makes things very complicated.

Fortunately, the forces of constraints often do no work, and then, we can model the constraints by a submanifold $N \subseteq N = \mathbb{R}^{3N}$.

Example 4.5. For example, the configuration of a rigid body is determined by the configuration of a collection of three points of the body not lying on the same line. Hence, we should start out with the configuration manifold $M = \mathbb{R}^9$. We demand that the body is rigid. Hence the distance between the different points should be constant. The constraints will then be given by

$$||x_i - x_j||^2 = \text{constant},$$

where x_k is the position of the k-th point ($k = 1, 2, 3$).

This suggests that a system consisting of m rigid bodies should be embedded into $M = (\mathbb{R}^9)^m$.

The dynamics in $M = \mathbb{R}^{3N}$, without is determined by a Lagrangian function $\mathcal{L} = T + U$, where U is the potential function modeling the interaction between the particles and T is the sum of the kinetic energies of the N particles, ie,

$$T = \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2,$$

where v_i is the velocity of particle nr. i. Hence, $(\mathbb{R}^{3N}, \mathcal{L})$ is a simple mechanical system. From the above discussion, the restriction of this system to a submanifold N is a simple mechanical system.

Often it is convenient to determine N independent of the $\mathbb{R}^{3N}$-model. In general we must however determine the Lagrangian from a more or less explicit (local) embedding of N into $\mathbb{R}^{3N}$

Example 4.6. The kinematics of a rigid body is determined by three orthogonal principal axes p_1, p_2, p_3 (and the corresponding moments of inertia) and the position of the center of mass. Hence, the configuration space N can be described as $\mathbb{R}^3 \times SO(3)$. If we are to determine the proper kinematic geometry of N, we need to an argument closely related to an argument where we embed N into $\mathbb{R}^9$.

5. Symmetries and conservation laws

Here we will consider Lie groups acting smoothly on configuration manifolds of Lagrangian systems and the conservation laws coming from such actions.

5.1. Symmetries. On any manifold M, we can consider the group $\mathrm{Diff}(M)$ consisting of diffeomorphisms on M acting in the natural way. This group also acts naturally on the tangent space TM by

$$\mathrm{Diff}(M) \times TM \to TM : \quad (\varphi, v_p) \mapsto T\varphi v_p.$$

Hence, we also get an action

$$\mathrm{Diff}(M) \times \mathcal{F}(TM) \to \mathcal{F}(TM) : \quad (\varphi, L) \mapsto \varphi^*(L), \tag{1}$$

where $\varphi^*(L)(v) = L(T\varphi(v))$ for all $v \in TM$.

Now, assume that $(M, \mathcal{L})$ is a Lagrangian system. Associated with this system, we have the set of diffeomorphisms of the configuration space M for which $\mathcal{L}$ is invariant, i,e, the isotropy group

$$\mathrm{Diff}^L(M) = \{\varphi : M \approx M : \varphi^* L = L\} \subseteq \mathrm{Diff}(M).$$

This group will be called the group of Lagrangian symmetries of $(M, \mathcal{L})$, and is denoted by $\mathrm{Iso}(M, \mathcal{L})$.

Now we assume that $(M, \mathcal{L})$ is a simple mechanical system with kinematic metric $\mathfrak{m}$ and potential function U. We can consider the quadratic form Q associated with the metric $\mathfrak{m}$. This gives us a function $Q : TM \to \mathbb{R}$. U can also be regarded as a function on TM, sending the tangent vector $v_p \in T_pM$ to the value $U(p) \in \mathbb{R}$. This gives us two more isotropy subgroups of $\mathrm{Diff}(M)$:

$$\mathrm{Iso}(M, K) = \mathrm{Diff}^Q(M) = \{\varphi : M \to M | \varphi^* Q = Q\}$$
$$\mathrm{Diff}^U(M) = \{\varphi : M \to M | \varphi^* U = U\} = \{\varphi : M \to M | U(\varphi(p)) = U(p) \forall p \in M\}.$$

In this situation, we get the following result

Proposition 5.2. *The group of Lagrangian symmetries of a natural Lagrangian system* $(M, \mathcal{L})$ *,*

$$\mathrm{Iso}(M, \mathcal{L}) = \mathrm{Iso}(M, K) \cap \mathrm{Diff}^U(M)$$

as a subgroup of $\mathrm{Diff}(M)$.

That is: A diffeomorphism $\varphi : M \to M$ *is a Lagrangian symmetry if and only if* φ *is a kinematic isometry preserving the potential function* U.

Proof. Expressing the Lagrange function in terms of $\mathfrak{m}$ and U, we see that $\varphi \in \mathrm{Iso}(M, \mathcal{L})$ if and only if

$$\mathfrak{m}(T\varphi v_p, T\varphi v_p) + U(\varphi(p)) = \mathfrak{m}(v_p, v_p) + U(p) \tag{3}$$

for all $p \in M$ and all $v_p \in T_pM$.

Now, assume that $\varphi \in \mathrm{Iso}(M, \mathcal{L})$. Using the equation above on the 0-section on the tangent bundle, we infer that

$$U(\varphi(p)) = U(p)$$

for all $p \in M$. Hence is $\varphi \in \mathrm{Diff}^U(M)$. Now, we can consider the function $\mathcal{L} - U$. This will obviously be φ-invariant, since both $\mathcal{L}$ and U are,, and hence is

$$\mathfrak{m}(T\varphi v_p, T\varphi v_p) = \mathfrak{m}(T\varphi v_p, T\varphi v_p) + U(\varphi(p)) - U(p) = \mathfrak{m}(v_p, v_p) + U(p) - U(p) = \mathfrak{m}(v_p, v_p).$$

This implies that $\varphi \in \mathrm{Iso}(M, K) \cap \mathrm{Diff}^U(M)$

Conversely we assume that $\varphi \in \mathrm{Iso}(M, \mathfrak{m}) \cap \mathrm{Diff}^U(M)$. Then we immediately see from the condition (3) that $\varphi \in \mathrm{Iso}(M, \mathcal{L})$.

This proves that $\mathrm{Iso}(M, \mathcal{L}) = \mathrm{Iso}(M, \mathfrak{m}) \cap \mathrm{Diff}^U(M)$ □

Remark 5.4. In the case that we replace the potential function U by a 1 form F modeling a force field, we replace $\mathrm{Diff}^U(M)$ with $\mathrm{Diff}^F(M)$, the group consisting of diffeomorphisms φ such that $\varphi^* F = F$. In this case, we will *define* the group $\mathrm{Iso}(M, \mathfrak{m}, F)$ of *symmetries of the simple mechanical system* $(M, \mathfrak{m}, F)$ by

$$\mathrm{Iso}(M, \mathfrak{m}, F) = \mathrm{Iso}(M, \mathfrak{m}) \cap \mathrm{Diff}^F(M)$$

5.2. The momentum map. Associated with a Lie group G, there is a Lie algebra $\mathfrak{g}$ canonically identified with the tangent space T_eG at the identity $e \in G$.

Now we will consider a Lie group G acting on a natural Lagrangian system $(M, \mathcal{L})$ by Lagrangian symmetries. This action can be regarded in a lot of equivalent ways:

First of all, we have the group homomorphism $\varphi : G \to \mathrm{Iso}(M, \mathcal{L})$ sending the group element $g \in G$ to a diffeomorphism $\varphi_g : M \to M$.

Also, given any $m \in M$, we get a smooth map $\psi_m : G \to M$, given by $\psi_m(g) = \varphi_g(m)$. Hence, we have a map

$$\psi : M \to C^\infty(G, M) \quad m \in M \mapsto \psi_m.$$

This map induces the tangent map

$$T_e\psi_m : \mathfrak{g} = T_eG \to T_mM,$$

giving the vector bundle map

$$\begin{array}{ccc} \mathfrak{g} \times M & \xrightarrow{\psi_*} & TM \\ {\scriptstyle\pi}\downarrow & & {\scriptstyle\pi}\downarrow \\ M & \xrightarrow{1_M} & M, \end{array} \tag{5}$$

which induces the dual map

$$\begin{array}{ccc} \mathfrak{g}^* \times M & \xleftarrow{\psi^*} & T^*M \\ {\scriptstyle\pi}\downarrow & & {\scriptstyle\pi}\downarrow \\ M & \xleftarrow{1_M} & M. \end{array}$$

Together with the inertia operator $\tilde{\mathfrak{m}}$ and the projection $M \times \mathfrak{g}^* \to \mathfrak{g}^*$, this defines the G*-momentum map* $\mathcal{J}_G$ as the composition

$$TM \xrightarrow{I} T^*M \xrightarrow{\psi^*} \mathfrak{g}^* \times M \xrightarrow{pr_2} \mathfrak{g}^*. \tag{6}$$

Given $v_p \in T_pM$ and $\xi \in \mathfrak{g}$,

$$\mathcal{J}_G(v_p)(\xi) = K_p(v_p, X_\xi), \tag{7}$$

where X_ξ is the vector field $x \mapsto \psi_*(\xi, x)$

Exercise 1. For two conjugate subgroups $H, K \subseteq G$, explore the relationship between $\mathcal{J}_H$ and $\mathcal{J}_K$, and discover the generalized Steiner's theorem relating angular momenta about different axes.

5.3. Noether's theorem. Now we will connect group actions and conservation laws via momentum maps.

5.3.1. *Some helpful results.* First we need a lemma concerning Killing fields. Recall that a Killing field X on a Riemannian manifold (M, K) is a vector field whose flow consists of isometries. In terms of the Lie derivative $\mathfrak{L}_X$, this is equivalent to $\mathfrak{L}_X K = 0$.

Lemma 5.8. (Characterisation of Killing fields). *Suppose that (M, K) is a Riemannian manifold with Riemannian connection ∇. A vector field $X \in \mathfrak{X}(M)$ is a Killing field if and only if*

$$\mathfrak{m}(\nabla_Y X, Z) + \mathfrak{m}(\nabla_Z X, Y) = 0$$

for all $Y, Z \in \mathfrak{X}(M)$.

Proof. Recall form section 1.1 on page 5 that $\mathfrak{L}_X$ commutes with contractions. sinve $\mathfrak{L}_X\mathfrak{m} = 0$ and ∇ is symmetric,

$$\begin{aligned} X(\mathfrak{m}(Y, Z)) &= \mathfrak{m}(\mathfrak{L}_X Y, Z) + \mathfrak{m}(Y, \mathfrak{L}_X Z) \\ &= \mathfrak{m}(\nabla_X Y, Z) - \mathfrak{m}(\nabla_Y, X) + \mathfrak{m}(Y, \nabla_X Z) - \mathfrak{m}(Y, \nabla_Z X). \end{aligned}$$

However, since ∇ is compatible with the metric $\mathfrak{m}$,

$$X(\mathfrak{m}(Y,Z)) = \mathfrak{m}(\nabla_X Y, Z) + \mathfrak{m}(Y, \nabla_X Z),$$

and the result follows from subtraction of the two equations. □

Another result which is far more obvious is the following:

LEMMA 5.9. (Invariance of functions). *Given a smooth manifold M and elements $f \in \mathcal{F}(M)$, and $X \in \mathfrak{X}(M)$. Then f is invariant under the flow generated by X if and only if $X(f) = 0$*

This shows that vector fields X generating symmetries of the simple mechanical system $(M, \mathfrak{m}, U)$ are characterized by $\mathfrak{L}_X \mathfrak{m} = 0$, $\mathfrak{L}_X U = XU = 0$. Since $X \mapsto \mathfrak{L}_X$ is a Lie algebra homomorphism it is easy to justify that we have Lie algebras

$$\mathfrak{X}^{\mathfrak{m}}(M) = \{X \in \mathfrak{X} | \mathfrak{L}_X \mathfrak{m} = 0\}$$
$$\mathfrak{X}^{U}(M) = \{X \in \mathfrak{X} | \langle X, U\rangle = 0\}$$
$$\mathfrak{X}^{\mathcal{L}}(M) = \{X \in \mathfrak{X} | \mathfrak{L}_X \mathcal{L} = 0\} = \mathfrak{X}^{\mathfrak{m}}(M) \cap \mathfrak{X}^{U}(M),$$

where we use the notation

$$\mathfrak{L}_X L(v) = \frac{d}{dt}|_{t=0}(\mathcal{L}(T\theta_t v)),$$

where θ_t denotes the local flow generated by X.

5.3.2. *Noether's theorem.* Now we are able to prove a theorem linking together symmetry group actions an conservation laws.

THEOREM 5.10. (Noether's theorem). *Assume that $(M, \mathcal{L})$ is a natural Lagrangian system on which a Lie group G acts by Lagrangian symmetries. Let $\gamma : [a,b] \to M$ be a motion of $(M, \mathcal{L})$. In this situation, the momentum map $\mathcal{J}_G$ is constant along γ. That is*

$$\mathcal{J}_G(\dot\gamma(t)) = \mathcal{J}_G(\dot\gamma(a)) \in \mathfrak{g}^* \quad \textit{for all } t \in [a,b].$$

PROOF. Let $\mathfrak{m}$ be the kinematic metric and ∇ be the kinematic connection associated with $(M, \mathcal{L})$, and we denote by U the potential function, and recall that G acts by $\mathfrak{m}$-isometries leaving U invariant.

Now we let $\xi \in \mathfrak{g}$ be a fixed element. We will show that $\langle \mathcal{J}_G(\dot\gamma(t)), \xi\rangle$ is a constant along γ. We recall equation (7), and see that

$$\langle \mathcal{J}_G(\dot\gamma(t)), \xi\rangle = K_{\gamma(t)}(\dot\gamma(t), X_\xi(\gamma(t)),$$

where $X_\xi : p \mapsto \psi_*(\xi, p)$ is the Killing field associated with ξ, and that $X_\xi(U) = 0$. The time derivative of this is equal to

$$\frac{d}{dt} K_{\gamma(t)}(\dot\gamma(t), X_\xi) = K_{\gamma(t)}(\nabla_{\dot\gamma}\dot\gamma(t), X_\xi) + K_{\gamma(t)}(\dot\gamma(t), \nabla_{\dot\gamma} X_\xi). \tag{11}$$

The first term on the right hand side is equal to $K_{\gamma(t)}(\blacktriangledown U, X_\xi)$ by Newton's equation. But, this is equal to 0, since U is assumed to be G-invariant. The second term is equal to 0 by lemma 5.8. Hence is the left side of equation (11) equal to 0, and we conclude that

$$\langle \mathcal{J}_G(\dot\gamma(t)), \xi\rangle = K_{\gamma(t)}(\dot\gamma(t), X_\xi)$$

is constant along motions γ of the system. Since we proved this for an arbitrary $\xi \in \mathfrak{g}$, this implies that $\mathcal{J}_G(\dot\gamma(t))$ is a constant in $\mathfrak{g}^*$. □

Remark 5.12. This theorem has a $(M, \mathfrak{m}, F)$-counterpart. Using equation (11), we must replace the first term on the right side by $\langle F_{\gamma(t)}, X_\xi \rangle$, and Noether's theorem comes on the form

$$\frac{d}{dt}\mathcal{J}_G(\dot{\gamma}(t)) = \psi^*_{\gamma(t)}F,$$

where $\psi^*_x F$ is the element of $\mathfrak{g}*$ such that $\langle \psi^*_x F, \xi \rangle = \langle F_x, \psi_x \xi \rangle = \langle F_x, X_\xi \rangle$.

This formula applies for all choices of F as long as G acts on $(M, \mathfrak{m})$ by isometries. Hence, we can use it even if F is not G-invariant.

Remark 5.13. Note that the substance of the proof of Noether's theorem is contained in lemma 5.8. The rest of the proof is a simple computation.

Remark 5.14. This theorem may also be formulated at the level of infinitesimal symmetries. A killing field such that $X(U) = 0$ is called an inifinitesimal symmetry of the simple mechanical system. From the proof of Noether's theorem, we realize that the quantity $\mathfrak{m}(X, \dot{\gamma})$ is conserved along motions of the Lagrangian system. Hence, the conservation law is properly represented by the 1-form $\tilde{\mathfrak{m}}X$

Conversely, if we start with a conservation law given by a 1-form ω such that $\omega(\dot{\gamma})$ is conserved along Lagrangian motions, then the vector field $X = \tilde{\mathfrak{m}}^{-1}\omega$ will be an infinitesimal symmetry of the Lagrangian system. Hence, there is a $1-1$-correspondance between symmetries and conservation laws represented by 1-forms.

Larger collections of infinitesimal symmetries can also be represented by Lie algebra (anti)-homomorphisms $\mathfrak{g} \to \mathfrak{X}(()M)$, such as in the case of a smooth action of a Lie group. In this general setting, we can define a momentum map $J : TM \to \mathfrak{g}^*$ in the same way as above. This can be regarded as an $\mathfrak{g}^*$-valued 1-form on M. The momentum map-construction gives a $1-1$-correspondance between the set of Lie algebras (anti)-homomorphisms $\mathfrak{g} \to \mathfrak{X}(M)$ going into the set of infinitesimal symmetries and the set of $\mathfrak{g}^*$-valued 1-forms giving constants of motion.

5.4. Equivariance of the momentum map. We prove that the momentum map is equivariant with respect to the coadjoint action on $\mathfrak{g}^*$.

As stated above, the action φ of G on M is lifted in a natural way to an action $T\varphi$ of G on the tangent bundle TM, where $T\varphi(g, v_p) = T_p\varphi_g(v_p)$.

We also have a natural action of G on $\mathfrak{g} \times M$ given by the adjoint action

$$Ad : G \times \mathfrak{g}M \to \mathfrak{g}M : \quad (g, (m, \xi)) \mapsto (gm, Ad_g \xi)$$

Now I want to prove that the map ψ_* in (5) is equivariant with respect to those actions. But, first, we need to look at this on a lower level.

Given any $g \in G$ and $m \in B$ we obviously have a commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{\psi_m} & M \\ {\scriptstyle \mathrm{Ad}_g}\big\downarrow & & {\scriptstyle \varphi_g}\big\downarrow \\ G & \xrightarrow{\psi_{gm}} & M, \end{array}$$

since $\varphi_g \circ \varphi_m(h) = \varphi_{gh}m = \varphi_{(ghg^{-1})g}m = \psi_m \circ Ad_g(h)$ for all $h \in G$. At $e \in G$, this induces the commutative diagram

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\psi_{m*}} & T_mM \\ {\scriptstyle \mathrm{Ad}_g}\big\downarrow & & {\scriptstyle \varphi_{g*}}\big\downarrow \\ \mathfrak{g} & \xrightarrow{\psi_{gm*}} & T_{gm}M \end{array}, \tag{15}$$

which proves that we get a commutative diagram

$$\begin{array}{ccc} G \times (\mathfrak{g} \times M) & \xrightarrow{Ad} & \mathfrak{g} \times M \\ {\scriptstyle 1_G \times \psi_*}\big\downarrow & & {\scriptstyle \psi_*}\big\downarrow \\ G \times TM & \xrightarrow{T\varphi} & TM \end{array}$$

since

$$\begin{aligned} \psi_* \circ Ad(g,(m,\xi)) &= \psi_{gm*} \circ Ad_g(\xi) \\ &= \varphi_{g*} \circ \psi_{m*}(\xi) \\ &= T\varphi \circ (1_G \times \psi_*)(g,(m,\xi)) \end{aligned}$$

This proves that ψ_* is G-equivariant.

If we take the dual of the diagram (15), we get the diagram

$$\begin{array}{ccc} \mathfrak{g}^* & \xleftarrow{\psi_m^*} & T_m^*M \\ {\scriptstyle \mathrm{Ad}_g^*}\big\uparrow & & {\scriptstyle \varphi_g^*}\big\uparrow \\ \mathfrak{g}^* & \xleftarrow{\psi_{gm}^*} & T_{gm}^*M. \end{array}$$

and in a similar way as above we prove that the diagram

$$\begin{array}{ccc} G \times (\mathfrak{g}^* \times M) & \xleftarrow{Ad^*} & \mathfrak{g}^*M \\ {\scriptstyle 1_G \times \psi^*}\big\uparrow & & {\scriptstyle \psi^*}\big\uparrow \\ G \times T^*M & \xleftarrow{\varphi^*} & T^*M \end{array}$$

is commutative, proving that φ^* is G-equivariant with respect to the right coadjoint action Ad^* on $\mathfrak{g}^*M$ and the action φ^* of G on T^*M given by $\varphi_g^*(\omega_m)(v_{g^{-1}m}) = \omega_m(\varphi_{g*}v_{g^{-1}m})$. Note that both Ad^* and φ^* are right actions.

The inertia operator $I : TM \to T^*M$ is also in a sense G-equivariant. We have the usual action

$$T\varphi : G \times TM \to TM,$$

and also the left action

$$\tilde{\varphi}^* : G \times T^*M \to T^*M$$

given by $\tilde{\varphi}^*(g,\omega) = \varphi^*(g^{-1},\omega)$. Here we see that

$$\varphi_{g^{-1}}^*(I(v_m))(w_{gm}) = \mathfrak{m}(v_m, \varphi_{g^{-1}*}w_{gm}) = \mathfrak{m}(\varphi_{g*}v_m, w_{gm}) = I(\varphi_{g*}v_m)(w_{gm})$$

And hence is the diagram

$$\begin{array}{ccc} G \times TM & \xrightarrow{T\varphi} & TM \\ {\scriptstyle 1_G \times I}\big\downarrow & & {\scriptstyle I}\big\downarrow \\ G \times T^*M & \xrightarrow{\tilde{\varphi}^*} & T^*M \end{array}$$

commutative.

Now, if we define the (left) coadjoint action $\tilde{\mathrm{Ad}}^* : G \times \mathfrak{g}^*M \to \mathfrak{g}^*M$ by $\tilde{\mathrm{Ad}}^*(g,(m,\xi)) = \mathrm{Ad}^*(g^{-1},(m,\xi))$ we see that we get a commutative diagram

$$\begin{array}{ccccccc} G\times TM & \xrightarrow{1_G\times I} & G\times T^*M & \xrightarrow{1_G\times\psi^*} & G\times(\mathfrak{g}^*\times M) & \xrightarrow{1_G\times pr_1} & G\times\mathfrak{g}^* \\ {\scriptstyle T\varphi}\downarrow & & {\scriptstyle\tilde{\varphi}^*}\downarrow & & {\scriptstyle\tilde{\mathrm{Ad}}^*}\downarrow & & {\scriptstyle\tilde{\mathrm{Ad}}^*}\downarrow \\ TM & \xrightarrow{I} & T^*M & \xrightarrow{\psi^*} & \mathfrak{g}^*\times M & \xrightarrow{pr_2} & \mathfrak{g}^*, \end{array}$$

where the last vertical arrow denotes the ordinary coadjoint action of G on $\mathfrak{g}^*$ given by

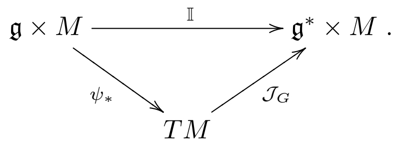
$$\langle \tilde{\mathrm{Ad}}^*_g\mu, \xi\rangle = \langle \mu, Ad_{g^{-1}}\xi\rangle.$$

This proves the following result.

PROPOSITION 5.16. *The G-momentum map $\mathcal{J}_G$ is a G-equivariant map $TM \to \mathfrak{g}^*$ with respect to the coadjoint action of G on $\mathfrak{g}^*$ and the lifted action of G on TM.*

5.5. A right inverse of the momentum map. Assume that we are given a fixed value $\mu \in \mathfrak{g}^*$ of the momentum map $\mathcal{J}_G$. For a point $x \in M$, this value μ may or may not be in the image of T_xM under the G-momentum map $\mathcal{J}_G$. The best that we can hope for is that there is a submanifold $M_\mu \subseteq M$ such that $\mu \in \mathrm{im}(\mathcal{J}_G|_x)$ for all $x \in M_\mu$. First we do a general observation:

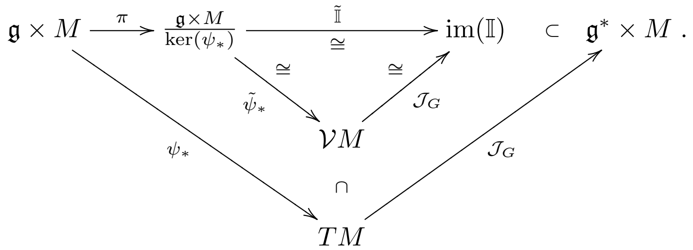
We have the following commutative triangle:



To see that this is commutative, we compute the effect on an element $\eta \in \mathfrak{g}$: Let $(\xi, m) \in \mathfrak{g}\times M$. Then $\langle \tilde{\mathfrak{m}}(\xi,m),\eta\rangle = \mathfrak{m}_m(\psi_m(\xi),\psi_m(\eta) = \langle \mathcal{J}_G(\psi_m(\xi)),\eta\rangle = \langle \mathcal{J}_G\circ\psi(\xi,m),\eta\rangle$.

By the assumption that all the isotropy subgroups are conjugate, ψ_* is a bundle map of constant rank, and hence $\ker(\psi_*) \subseteq \mathfrak{g}\times M$ is a subbundle. This bundle is i fact identical to the subset $\ker\mathbb{I} \subseteq \mathfrak{g}\times M$: $\psi(\xi,m) = 0$ if and only if $\mathfrak{m}_m(X_xi, v) = 0$ for all vectors $v \in T_mM$. But X_ξ is vertical, and hence this condition holds if and only if it holds for every vertical vector $v \in \mathcal{V}_mM$. But $\psi : \mathfrak{g} \to \mathcal{V}_mM$ is surjective, and hence the condition holds if and only if $\mathfrak{m}_m(X_xi, X_\eta) = 0$ for all $\eta \in \mathfrak{g}$.. But, the last condition is just another form of $\langle \mathbb{I}(\xi,m),\eta\rangle = 0$. Hence $\psi_*(\xi,m) = 0$ if and only if $\mathbb{I}(\xi,m) = 0$.

Because the image of ψ_* is equal to $\mathcal{V}M$, this gives us the inner triangle of isomorphisms



This has a useful application: For a given μ, we consider the set M_μ defined above. We denote by $\mathcal{V}M \downarrow M_\mu$ the restriction of the vertical bundle $\mathcal{V}M$ to M_μ. The momentum map $\mathcal{J}_G : \mathcal{V}M \downarrow M\mu \to \mathrm{im}(\mathbb{I}) \downarrow M_\mu$ is an isomorphism. Hence, the section (μ, m) in $\mathrm{im}(\mathbb{I}) \downarrow M_\mu$ corresponds to a unique section X_μ in $\mathcal{V}M \downarrow M_\mu$. This vector field can be characterized in the following way: $X_\mu(m)$ is the unique vertical vector such that $\mathcal{J}_G(X_\mu(m)) = \mu$.

A motion γ in M with $\mathcal{J}_G(\dot\gamma) = \mu$ must be a curve in M_μ. further more, the vertical component of the velocity vector $\mathrm{pr}_{\mathcal{V}}\dot\gamma(t) = X_\mu(\gamma(t)$, since $X_\mu(\gamma(t))$ is the unique vertical vector with $\mathcal{J}_G(X_\mu(m)) = \mu$.

Then, we should be interested in the differential topology of M_μ. I cannot say anything general about this. Fortunately, it often happens that M_μ is a submanifold of M. Hence, we will assume this whenever we talk about M_μ even though it is not generally true.

However, it is true in one very important situation: If G acts freely, then it is easy to see that $\psi_* : \mathfrak{g} \times M \to TM$ is injective. Following from this, $\mathbb{I}$ is injective, and of dimensional reasons, it is an isomorphism. Hence $\psi_* \circ \mathbb{I}^{-1}$ is a right inverse of $\mathcal{J}_G$, so that, for any $m \in M$ and $\mu \in \mathfrak{g}^*$ there is a $v \in \mathcal{V}_m M$, namely $v = \psi_* \circ \mathbb{I}^{-1}(\mu, m)$, such that $\mathcal{J}_G(v) = \mu$. We give another example:

EXAMPLE 5.17. We let $M = \mathbb{R}^3$, $G = SO(3)$ and ψ the standard representation. We use the standard identifications $\mathfrak{g} = \mathbb{R}^3$, $\mathfrak{g} = \mathbb{R}^3$, so that $[\xi, \eta] = \xi \times \eta$, and if $\xi = (x_1, x_2, x_3)$ and $\mu = (m^1, m^2, m^3)$, then $\langle \xi, \mu \rangle = \Sigma_i a_i m^i$. We equip $\mathbb{R}^3$ with the usual Riemannian metric. Ie, we consider the motion of a particle in $\mathbb{R}^3$ with mass $m = 1$. In this setting, the $SO(3)$-momentum is the usual angular momentum with respect to the origin 0.

For a motion γ, the $SO(3)$-momentum is given by $\mu(t) = \gamma \times \dot\gamma$. But, bu the well known properties of $\times$, $\mu(t) \perp \gamma(t)$ and hence $M_\mu \subseteq \mu^\perp$, the subspace of $\mathbb{R}^3$ perpendicular to μ. Conversely, if $x \in \mu^\perp$, then there is an $y \in span\{x \times \mu\}$ such that $x \times y = \mu$. Hence, $M_\mu = \mu^\perp$. This is clearly a submanifold of $\mathbb{R}^3$.

EXERCISE 2. Make an example where $M_\mu \subseteq M$ is a horrible subspace.

CHAPTER 2

Reduction theory

1. Heuristic description of a method of reduction

The next sections contains some ideas about how we would like configuration space reduction to work. The objective is to understand how we can use configuration space symmetries to give a simpler formulation of the equations of motion, and to get a better understanding of the dynamics.

Motions γ of the mechanical system on the configuration space M gives us curves $\delta = \pi \circ \gamma$ in the reduced configuration space $N = M/G$. We hope to be able to describe the reduced curves δ as motions of a mechanical system on the reduced configuration space N. Together with this, we want a simple method to reconstruct γ from δ given that we know some appropriate initial conditions. It turns out that this is possible only in some very special cases. Hence, our main task will be to try to understand the difficulties with this approach.

First we will describe one idea about how this should work in some particular nice situations, and simultaneously discover the obstructions against usage of such a method in a general setting. However, we will always assume that we have the optimal differential topological situation, and thus concentrate on the more geometric obstructions. For instance, we assume that the G-action on the configuration manifold M gives us a smooth submersion $\pi : M \to N = M/G$ onto the reduced configuration manifold.

The idea behind this reduction method is that if we know the G-momentum μ of the motion γ, we are able to determine the component $\mathrm{pr}_{\mathcal{V}}\dot{\gamma}$ of $\dot{\gamma}$ along G-orbits at any relevant point. Referring to section 5.5, page 23, we write $\mathrm{pr}_{\mathcal{V}}\dot{\gamma} = X_\mu$. Hence, if we know the horizontal component $\mathrm{pr}_{\mathcal{H}}\dot{\gamma}$ of $\dot{\gamma}$, we will know $\dot{\gamma}$, and we canreconstructt γ. But, $\mathrm{pr}_{\mathcal{H}}\dot{\gamma} = \mathrm{L}_0\dot{\delta}$, where $\delta = \pi \circ \gamma$. Hence, if we know δ, we know $\mathrm{pr}_{\mathcal{H}}\dot{\gamma}$.

2. Systems admitting free actions

Now we consider a differential-topological optimal simple mechanical system $(M, \mathfrak{m}, U)$ with symmetry group G, acting properly on M. For the moment, we assume that G acts freely.

Let $\mu \in \mathfrak{g}$ be a given momentum. There is a unique vector field X_μ on M with $\mathcal{J}_G(X_\mu(m)) = \mu$ for all $m \in M$. From the discussion above, we see that we now know the vertical component of the velocity of every motion with G-momentum μ.

In this situation, we may decompose the kinetic energy of a motion into two parts,

$$T = T^{\mathcal{V}} + T^{\mathcal{H}},$$

where

$$T^{\mathcal{V}} = \frac{1}{2}\mathfrak{m}(\mathrm{pr}_{\mathcal{V}}\dot{\gamma})^2 = \frac{1}{2}\mathfrak{m}X_\mu^2 \quad \text{and} \quad T^{\mathcal{H}} = \mathfrak{m}(\mathrm{pr}_{\mathcal{H}}\dot{\gamma})^2 = \mathfrak{m}_N\dot{\delta}^2.$$

This gives us the following expressions for the Lagrangian function:

$$\mathcal{L} = T + U = (T^{\mathcal{H}} + T^{\mathcal{V}}) + U = T^{\mathcal{H}} + (T^{\mathcal{V}} + U) = T^{\mathcal{H}} + \tilde{U}$$

$T^{\mathcal{H}}$ can be regarded as a function on the tangent bundle TN of the reduced configuration space. But, why should not $\tilde{U}$ be a function on N. Obviously this is not generally true. However, let us imagine that this is true, and look at the variational principle for the Lagrangian systems involved. We write $\tilde{\mathcal{L}}_\mu = T^{\mathcal{H}} + \tilde{U} = \tilde{T} + \tilde{U}$ for the Lagrangian down on the reduced configuration space. Since $2T^{\mathcal{H}}(\dot{\gamma}) = \mathfrak{m}(\mathrm{pr}_{\mathcal{H}}\dot{\gamma})^2 = \mathfrak{m}_N\delta^2$, the Lagrangian system $(N, \tilde{\mathcal{L}}_\mu)$ can be regarded as a simple mechanical system with kinematic metric $\mathfrak{m}_N$ and potential function $\tilde{U}$

2.1. The variational approach. First, we assume that γ is a motion of the simple mechanical system $(M, \mathfrak{m}, U)$. This implies that the integral functional

$$\Lambda[\gamma] = \int_a^b [T(\dot{\gamma}) + U(\gamma)]dt$$

is stationary at γ with respect to smooth variations keeping the endpoints fixed. But, this implies that Γ is stationary at γ with respect to variations that keep *both* endpoints *and* momentum μ fixed. Let $\gamma_s(t)$ be such a variation. Such a variation gives us a variation $\delta_s(t)$ of the reduced motion δ. Now,

$$\Gamma[\gamma] = \int_a^b \mathcal{L}(\dot{\gamma})dt = \int_a^b \tilde{\mathcal{L}}(\dot{\delta})dt =^{def} \tilde{\Lambda}[\delta],$$

and hence, $\tilde{\Lambda}$ is stationary at δ with respect to variations of δ that are projections of variations of γ keeping endpoints and momentum fixed.

However, given any variation δ_s of δ keeping endpoints fixed, we can construct a variation γ_s of γ keeping endpoints fixed, such that $\pi \circ \gamma_s = \delta_s$, simply by letting $s \mapsto \gamma_s(t_0)$ be the horizontal lifting of $s \mapsto \delta_s(t_0)$. Thus, δ must be a stationary point of $\tilde{\Lambda}$ with respect to smooth variations keeping endpoints fixed. Hence, the curve δ is a motion of the simple mechanical system $(N, \tilde{\mathcal{L}}_\mu)$, and the curve γ can be reconstructed from δ.

Now, we want to go in the reverse direction. Let $\delta : [0,1] \to N$ be a motion of the system $(N, \tilde{\mathcal{L}}_\mu)$. Given any $x \in \pi^{-1}(\delta(0))$, there is a unique motion $\gamma : [0, \varepsilon] \to M$ of $(M, \mathcal{L})$ with

$$\gamma(0) = x \quad \text{and} \quad \dot{\gamma}(0) = X_\mu(x) + \mathrm{L}_0\dot{\delta}(0).$$

We know that γ projects down to a motion $\tilde{\delta}$ of $(N, \tilde{\mathcal{L}}_\mu)$. Obviously, $\tilde{\delta}(0) = \delta(0)$ and $\dot{\tilde{\delta}}(0) = \dot{\delta}(0)$. Hence, by uniqueness of motions with given initial velocity, $\tilde{\delta} = \delta$ on their common domain of definition. Hence, δ is, up to correct choice of parameter domain, the projection of a motion of $(M, \mathcal{L})$. Hence, every motion of $(N, \tilde{\mathcal{L}}_\mu)$ is the projection of a motion in $(M, \mathcal{L})$.

From this discussion, we see that in the case of free group actions, we get into trouble when $\mathfrak{m}(X_\mu, X_\mu)$ is not G-invariant. Hence, we need another approach for that situation. We close this discussion by demonstration of some situations where this technique is effective:

2.2. Horizontal reduction. If we assume $\mu = 0$, the reduction is easy, since $\mu = 0$ is equivalent to $\mathrm{pr}_{\mathcal{V}}\dot\gamma = 0$. In this case, γ is equal to a horizontal lift $\mathrm{L}_0\delta$ of the reduced curve δ. Hence $\dot\gamma = \mathrm{L}_0\dot\delta$, and therefore

$$\nabla_{\dot\gamma}\dot\gamma = \nabla_{\mathrm{L}_0\dot\delta}\mathrm{L}_0\dot\delta = \mathrm{L}_0(\overset{N}{\nabla}_{\dot\delta}\,\dot\delta).$$

The potential function U on M projects to a function $\tilde U$ on N, where the gradients are related by

$$\blacktriangledown U = \mathrm{L}_0(\blacktriangledown \tilde U),$$

and hence

$$\nabla_{\dot\gamma}\dot\gamma = \blacktriangledown U \quad \Leftrightarrow \quad \overset{N}{\nabla}_{\dot\delta}\,\dot\delta = \overset{N}{\blacktriangledown}\,\tilde U.$$

We conclude that γ is a motion of $(M, \mathfrak{m}, U)$ with G momentum $\mu = 0$ if and only if γ projects to a motion δ of the Lagrangian system $(N, \mathfrak{m}_N, \tilde U)$.

Remark 2.1. This applies also to systems $(M, \mathfrak{m}, F)$ given by a generalized force F if the force F is both G-invariant and annihilates vertical vectors, ie, if the vector field $\tilde{\mathfrak{m}}^{-1}F$ is G-invariant and perpendicular to G-orbits.

2.3. Abelian Routh reduction. In this section, we assume that G is an abelian group acting freely[1] on M.

This is the typical situation when we have chosen a coordinate system $(q_1, \ldots, q_n)$ with one or more *cyclic* variables. That is variables (x_i) such that $\frac{\partial\mathcal{L}}{\partial x_i} = 0$. A collection of cyclic cordinates gives us a collection $\{X_i\}$ of coordinate vector fields, and as we all know, such vector fields commute, i.e, $[X_i, X_j] = 0$. Hence, the vector fields represents a commuttative (local) transformation group. A treatment of the reduction in this case is found in the article [**10**] written by E.J.Routh in 1877.

Since G is abelian, the adjoint and coadjoint representations are trivial, and hence, for a given momentum $\mu \in \mathfrak{g}^*$, the associated vector field $X_\mu \in \mathfrak{X}(M)$ is G-invariant. This follows from the formula $T\varphi_g X_\mu = X_{\mathrm{Ad}^*_g\mu} = X_\mu$, coming from the equivariance of the momentum map. See proposition 5.16 on page 22.

This implies that the function $\mathfrak{m}(X_\mu, X_\mu) \in \mathcal{F}(M)$ is G-invariant, since

$$\mathfrak{m}(X_\mu, X_\mu)(\varphi_g x) = \mathfrak{m}(T\varphi_g X_\mu, T\varphi_g X_\mu) = \mathfrak{m}(X_\mu, X_\mu)(x).$$

Because of this, there is a function R^μ on N such that $\mathrm{L}_0 R^\mu = \mathfrak{m}(X_\mu, X_\mu)$. The potential function U is also G-invariant, and projects to a function $\tilde U$ on N. We can write the Lagrangian function of the original system as

$$\mathcal{L} = T + U = T^{\mathcal{H}} + \mathrm{L}_0(R^\mu + \tilde U).$$

The function $T^{\mathcal{H}}$ depends only on the velocity of the projected curve δ. Hence, $T^{\mathcal{H}}(\dot\gamma) = \frac{1}{2}\mathfrak{m}_N\dot\delta^2$. Hence, there is a natural choice of a Lagrangian system on N, given by the Lagrange function

$$\tilde{\mathcal{L}}_\mu(v) = \frac{1}{2}\mathfrak{m}_N v^2 + (R^\mu + \tilde U)$$

As noted in section 2.1, the motions of $(M, \mathcal{L})$ are exactly the curves that projects to motions of $(N, \tilde{\mathcal{L}}_\mu)$.

[1] This is only slightly less general than assuming that the action has one unique isotropy type. $(G_x) = (G_y)$ implies $G_x = G_y$, since G is abelian. Hence, we can replace the action of G by a free action of G/G_x on M.

REMARK 2.2. This approach will be successful also when we only assume that μ is invariant under the coadjoint action of G on $\mathfrak{g}^*$. This can be found in for example Smale [**11**]. The horizontal reduction is a special case of this, since $0 \in \mathfrak{g}^*$ always is invariant under the coadjoint action of G on $\mathfrak{g}^*$.

2.4. Left-invariant mechaical system on a Lie group. Now, we let $M = G$ be a Lie group, and consider the left translation action of G on itself. We let $\mathcal{L} : TG \to \mathbb{R}$ be a G-symmetric Lagrangian function for a simple mechanical system on G. Since G acts transitively on itself, the potential function U must be a constant, and hence we may assume $U = 0$. Thus we see that $\mathcal{L}$ is completely determined by the kinematic metric $\mathfrak{m}$ which must be left invariant.

Now, a trivialisation of TG is given by

$$h_L : TG \to \mathfrak{g} \times G \quad h_L(v) = (TL_{\pi(v)^{-1}}\pi(v)),$$

the *left* trivialisation. Since $\mathfrak{m}$-is left invariant

$$\mathfrak{m}(v, w)(g) = \mathfrak{m}(TL_{g^{-1}}v, TL_{g^{-1}}w)(e), \quad \forall g \in G$$

and hence $\mathfrak{m}$ is determined by its action on tangent vectors at the identity element e, which is essentially the same as the elements of the Lie algebra $\mathfrak{g}$. Looking at this from the point of view of left invariant vector fields, we see that $\mathfrak{m}(X, Y)$ is a constant function if X, Y are left invariant.

To summarize this: A left invariant metric $\mathfrak{m}$ determines, and is determined by an inner product on the Lie algebra $\mathfrak{g}$ of left invariant vector fields. Now, we have a fixed inner product $(-, -)$ on $\mathfrak{g}$, and hence we also have a fixed isomorphism $\iota : \mathfrak{g} \to \mathfrak{g}^*$, induced by this inner product. The inner product $\mathfrak{m}$ is then expressed at $\mathfrak{g} \times G$ as

$$\mathfrak{m}((\xi, g), (\eta, g)) = (\xi, \eta)$$

By dualisation of the left trivialisation $h_L : TG \to \mathfrak{g} \times G$, we obtain a trivialisation $h_L^* : \mathfrak{g}^* \times G \to T^*G$. For $\mu \in \mathfrak{g}^*$, $g \in G$ and $v \in T_gG$, we have $h_L^*(\mu, g)(v) = \langle(\mu, g), (h_L v)\rangle = \langle \mu, TL_{g^{-1}}v\rangle$

Now we want to express the mass operator $m : TG \to T^*G$ with respect to the left trivialisations. Let $v_g, w_g \in T_gG$. $\langle m(v_g), w_g\rangle = \mathfrak{m}(v_g, w_g)$. Assume that $h_L(v_g) = (\xi, g), h_L(w_g) = (\eta, g)$. This means that $TL_{g^{-1}}v_g = \xi$, and so on. Since $\mathfrak{m}$ is left invariant, we have

$$\mathfrak{m}(v_g, w_g) = \mathfrak{m}(\xi, \eta) = (\xi, \eta) = \langle \iota(\xi), \eta\rangle = \langle \iota(\xi), TL_{g^{-1}w_g}\rangle.$$

But this expresses that $m(v_g) = h_L^*(\iota(\xi), g)$. Hence is the mass operator expressable as

$$m : \mathfrak{g} \times G \to \mathfrak{g}^* \times G, \quad m(\xi, g) = (\iota(\xi), g)$$

2.4.1. *The momentum map.* Associated with the action of G on itself and the kinematic metric $\mathfrak{m}$, there is a momentum map $\mathcal{J} : TG \to \mathfrak{g}^*$.

Let $g \in G$ be a fixed element. Then we have the mapping $\psi_g : G \to G : h \mapsto hg$, namely the right translation by g, R_g. The associated tangent map at e is then $T_eR_h : \mathfrak{g} \to T_gG$, this gives the bundle isomorphism $\psi : \mathfrak{g} \times G \to TG$ given by

$$\psi(\xi, g) = TR_g\xi$$

This is in fact a trivialisation of the tangent bundle, and we can call it the right trivialisation.

Now, we dualize this, to get a map $\psi^* : T^*G \to \mathfrak{g}^* \times G$. Explicitly,

$$\langle \psi^*(\omega_g), (\xi, g)\rangle = \omega_g \circ \psi(\xi, g) = \langle \omega_g, TR_g\xi\rangle$$

We can express this in terms of the left trivialisation of T^*G, and we get

$$\langle \psi^*(\mu, g), (\xi, g)\rangle = \langle \mu, TL_{g^{-1}}TR_g\xi\rangle = \langle \mu, Ad_{g^{-1}}\xi\rangle$$

This map is essentially the left coadjoint representation

$$Ad^* : \mathfrak{g}^* \times G \to \mathfrak{g}^* \times G \quad Ad^*(\mu, g) = (Ad^*_g\mu, g).$$

Form this, and the expression of the mass operator ι, we see that the momentum map $J : TG \to \mathfrak{g}^*$ defined by the compositions

$$\begin{array}{ccccccc} TG & \xrightarrow{m} & T^*G & \xrightarrow{\psi^*} & \mathfrak{g}^* \times G & \xrightarrow{\pi} & \mathfrak{g}^* \\ h_L \big\downarrow \cong & & h_L^* \big\uparrow \cong & & id \big\downarrow & & id \big\downarrow \\ \mathfrak{g} \times G & \xrightarrow{\iota} & \mathfrak{g}^* \times G & \xrightarrow{\psi^*} & \mathfrak{g}^* \times G & \xrightarrow{\pi} & \mathfrak{g}^* \end{array},$$

is given by

$$\langle J(\xi, g), \eta\rangle = \langle Ad^*_g\iota(\xi), \eta\rangle = \langle \iota(\xi), Ad_{g^{-1}}\eta\rangle = (\xi, Ad_{g^{-1}}\eta)$$

in terms of the left trivialisations. And hence,

$$J(\xi, g) = Ad^*_g\iota(\xi) \tag{3}$$

2.4.2. *The associated submersion.* The submersion associated with this symmetric system is $G \to *$, i.e, a rather trivial one. We have one vector field on $*$, namely the 0 section of the bundle $\{0\} \times * \to *$. The horizontal lifting of this vector field is the 0-section of the tangent bundle. We look at the non-horizontal motions:

2.4.3. *The vector field associated with a given momentum.* For a given $\mu \in \mathfrak{g}^*$, we get a unique vector field X_μ such that $J(X_{\mu_g}) = \mu$ for all $g \in G$. The simple reason of this is that the unprojected version

$$J : TG \to \mathfrak{g}^* \times G$$

of the momentum map is a bundle isomorphism.

First I want to compute X_μ in terms of the trivialisation $h_L : TG \cong \mathfrak{g} \times G$. Inverting equation (3), we see that the representative at $g \in G$ of X_μ, $\xi = \iota^{-1}Ad^*_{g^{-1}}(\mu)$. Hence is $(\xi, \eta) = \langle Ad^*_{g^{-1}}(\mu), \eta\rangle$ whenever $\eta \in \mathfrak{g}$. More important, $X_\mu(g) = TL_g(\iota^{-1}\mathrm{Ad}^*_{g^{-1}}(\mu))$.

2.4.4. *The equations of motion.* The equations of motion of the Lagrangian system $(G, \mathcal{L})$ will be

$$\nabla_{\dot\gamma}\dot\gamma = 0 \tag{4}$$

We know that the momentum map is constant along motions of this system. We also know that every initial state $v \in TG$ gives a unique initial momentum. Hence, from the existence and uniqueness-theorem of integral curves smooth vector fields, we infer that the motions of $(G, \mathcal{L})$ with initial angular momentum μ are exactly the integral curves of the vector field X_μ.

Hence, we have reduced the equation of motion to

$$\dot\gamma(t) = X_\mu(\gamma(t)) = T_eL_{\gamma(t)}(\iota^{-1}\mathrm{Ad}^*_{\gamma(t)^{-1}}(\mu)) \tag{5}$$

With this, we have reduced the equations of motion to a first order equation in an unusual way. We look at how this works in the case of the free rigid body:

EXAMPLE 2.6. The configuration space of the rigid body is the group $SO(3)$. For this group, we usually identify the Lie algebra $\mathfrak{g}$ with the space of skew symmetric 3×3-matrices. We may identify $\mathfrak{g}^*$ with the same space, and let $\mu \in \mathfrak{g}^*$ act on $\xi \in \mathfrak{g}$ by

$$\langle \xi, \mu \rangle = \mathrm{tr}(\mu\xi),$$

where $\xi\mu$ is the product of the matrices representing ξ and μ.

An inner product $(-,-)$ on $\mathfrak{g}$ may be represented by a symmetric matrix A, such that $(\xi,\eta) = \mathrm{tr}(\eta A \xi^T)$. In this case, the associated operator $\iota : \mathfrak{g} \to \mathfrak{g}^*$ is given by $\iota(\xi) = A\xi^T$, and the inverse is then given by $\iota^{-1}(\mu) = (A^{-1}\mu)^T$. For a motion $g(t)$ with $g^{-1}\dot{g}(t) = \xi(t)$, the momentum μ is equal to $g^{-1}A\xi^T g$.

The adjoint action of G on $\mathfrak{g}$ is given by $\mathrm{Ad}_g\xi = g\xi g^{-1}$ (Remember that both g and ξ are matrices), and hence

$$\langle \xi, \mathrm{Ad}_g^*\mu \rangle = \langle \mathrm{Ad}_g\xi, \mu \rangle = \mathrm{tr}(()\mu g\xi g^{-1}) = \mathrm{tr}(()g^{-1}\mu g \xi g^{-1} g) = \langle \xi, g^{-1}\mu g \rangle,$$

since $\mathrm{tr}(-)$ is Ad-invariant. We conclude that $\mathrm{Ad}_g^*\mu = g^{-1}\mu g$. Hence, we get equation (5) on the form

$$\dot{g} = g(A^{-1}g^{-1}\mu g)^T = gg^T\mu^T g(A^{-1})^T = \mu^T g(A^{-1})^T = -\mu g A^{-1}, \tag{7}$$

remembering that $g^{-1} = g^T$ when $g \in SO(3)$. If we, as above, write $\xi = g^T\dot{g}$, and we note that $\xi A = -A\xi$, since A is symmetric and ξ is anti-symemtric, we end up with

$$A\xi = g^T\mu g$$

Since, μ and A are constants, and $\mu = gA\xi g^T$, we get

$$A\dot{\xi} = \dot{g}^t\mu g + g^T\mu\dot{g} = \dot{g}^T gA\xi g^T g + g^T gA\xi g^T\dot{g} = \xi^T A\xi + A\xi\xi = [A\xi, \xi],$$

an equation that is called the Euler-equation. Finally, we deduce a generalization of this:

2.4.5. *The Euler-Poincaré equations on compact groups.* The reduction above is somewhat nontraditional, so I think that I will enclose the Euler-Poincare-equation. The essence is contained in the identity $\dot{\mu} = 0$, i.e., that the momentum is constant along motions of the Lagrangian system.

To be able to handle this situation, we must equip our Lie group G with a bi-invariant Riemannian metric $\langle\langle -,-\rangle\rangle$. The possibility of this is a well known fact. Taking $\langle\langle -,-\rangle\rangle$ for granted, we get an inner product $\langle\langle -,-\rangle\rangle$ on $\mathfrak{g}$ and a canonical isomorphism $\kappa : \mathfrak{g} \cong \mathfrak{g}^*$. $\langle\langle -,-\rangle\rangle$ is Ad-invariant, and the map $\kappa : \mathfrak{g} \to \mathfrak{g}^*$ will be $Ad - Ad^*$-equivariant, since

$$\langle \kappa Ad_g\xi, \eta \rangle = \langle\langle Ad_g\xi, \eta \rangle\rangle = \langle\langle \xi, Ad_{g^{-1}}\eta \rangle\rangle = \langle \kappa\xi, Ad_{g^{-1}}\eta \rangle = \langle Ad_g^*\kappa\xi, \eta \rangle$$

We can express $\mathfrak{m}(-,-)$ in terms of $\langle\langle -,-\rangle\rangle$: There exists an operator $A : \mathfrak{g} \to \mathfrak{g}$ such that $\mathfrak{m}(\xi,\eta) = \langle\langle A\xi, \eta \rangle\rangle$. Hence is $\iota(\xi) = \kappa(A\xi)$, and

$$J(\xi, g) = Ad_g^*\iota(\xi) = Ad_g^*\kappa A\xi = \kappa Ad_g A\xi$$

Using this, we translate $\dot{\mu} = 0$ into $\frac{d}{dt}Ad_{\gamma(t)}A\xi(t) = 0$, where $\xi(t) = TL_{\gamma(t)^{-1}}\dot{\gamma}(t)$, which is translated to

$$A\dot{\xi} = [A\xi, \xi].$$

EXAMPLE 2.8. When $G = SO(3)$, and we identify $\mathfrak{g}$ with $\mathbb{R}$ in the usual way, such that $[\xi, \eta] = \xi \times \eta$ (cross product of vectors), and we put A on the diagonal form $diag(\lambda_1, \lambda_2, \lambda_3)$, we get the Euler-equations

$$\lambda_1 \dot{x}_1 = (\lambda_2 - \lambda_3) x_2 x_3$$
$$\lambda_2 \dot{x}_2 = (\lambda_3 - \lambda_1) x_3 x_1$$
$$\lambda_3 \dot{x}_3 = (\lambda_1 - \lambda_2) x_1 x_2.$$

3. A possible general reduction procedure

A simple mechanical system $(M, \mathcal{L})$ can be regarded as a dynamical system on the manifold TM. Hence, for a Lagrangian symmetry action of a Lie group G on M we could expect to end up with a dynamical system on $(TM)/G$. For simplicity, we assume that G acts freely.

We can describe $(TM)/G$ as a vector bundle over M/G: We know that TM is canonical isomorphic to the Whitney sum $\mathcal{V}M \oplus \mathcal{H}M$, ie,the sum of the vertical distribution and the horizontal distribution coming from the Riemannian submersion $\pi : M \to M/G$. But both these bundles are G-bundles over M.

First we describe $(\mathcal{H}M)/G$: To vectors v, w in the horizontal distribution $\mathcal{H}M$ project to the same element of $(\mathcal{V}M)/G$ if and only if there is an element $g \in G$ with $T\varphi_g v = w$. But this is the case if and only if v and w are projected to the same vector in $T(M/G)$. Hence, $(\mathcal{H}M)/G \cong T(M/G)$.

Then we describe $(\mathcal{V}M)/G$: As noted implicitly in section 5.5, there is a bijection $\mathcal{H}M \cong \mathfrak{g} \times M$. $\mathcal{H}M$ is a G-bundle, and there is an action of G on $\mathfrak{g} \times M$ given by

$$(g, (\xi, m)) = (\mathrm{Ad}_g \xi, m)$$

From section 5.4 we see that the isomorphism $\mathcal{H}M \cong \mathfrak{g} \times M$ is in fact an isomorphism of G-bundles, and hence is $(\mathcal{H}M)/G \cong (\mathfrak{g} \times M)/G$. But $(\mathfrak{g} \times M)/G$ is a kind of a twisted product bundle with fibre $\mathfrak{g}$ over the base space M/G. Following Cendra et.al. [**3**], we use the notation $\tilde{\mathfrak{g}}$ for this bundle.

From this we see that we can describe $(TM)/G$ as the Whitney-sum

$$(TM)/G \cong T(M/G) \oplus \tilde{\mathfrak{g}}.$$

We denote the projection $TM \to T(M/G) \oplus \tilde{\mathfrak{g}}$ by π.

The Lagrangian function $\mathcal{L} : TM \to \mathbb{R}$ is, by assumption, G-invariant. Hence, there is a function $l : T(M/G) \oplus \tilde{\mathfrak{g}} \to \mathbb{R}$ such that $\mathcal{L} = l \circ \pi$.

This suggests that we could try to formulate a variational problem on $T(M/G) \oplus \tilde{\mathfrak{g}}$. The problem, as noted above is that it is difficult to know the proper restrictions on the variational principle. From the discussion in section 2.1, we see that we can allow all variations of projected curves $\delta = \pi \circ \gamma$. However, the correct choice of variations in the $\tilde{\mathfrak{g}}$-part is not obvious to me. This is treated by Marsden and Scheurle in [**8**], and they also state some explicite equations. In essence, I hope that this approach can lead to a dynamical system on $T(M/G) \oplus \tilde{\mathfrak{g}}$ consisting of a a set of equations strongly related to the Euler-Poincaré-equations on G, coupled to a set of equations coming from a simple mechanical system on M/G with a generalized force depending on the $\tilde{\mathfrak{g}}$-part of the system.

Bibliography

[1] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer Verlag, 2. edition, 1989.

[2] A. M. Bloch. *Nonholonomic Mechanics and Control*. Interdisciplinary Applied Mathematics. Springer, 2003.

[3] H. Cendra, J. E. Marsden, and T. S. Ratiu. *Lagrangian Reduction by stages*, volume 722. Memoirs of the American Mathematical Association, 2001.

[4] M. P. do Carmo. *Riemannian Geometry*. Birkhuser, 1992.

[5] A. D. Lewis F.Bullo. *Geometrical Control of Mechanical Systems*. Springer Verlag, 2005.

[6] Wu-Yi Hsiang and Eldar Straume. Kinematic geometry of triangles and the study of the three-body problem. www.arXiv.org, August 2006, to appear in Lobachevski Journal of Mathematics 2007.

[7] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume I. Interscience Publishers, 1. edition, 1963.

[8] J. E. Marsden and Jürgen Scheurle. The reduced euler-lagrange equations. *Fields Institute Communications*, (1):139–164, 1993.

[9] W. M. Oliva. *Geometric Mechanics*. Number 1798 in Lecture Notes in Mathematics. Springer, 2002.

[10] E. J. Routh. Stability of a given state of motion. In A. T. Fuller, editor, *Stability of motion*. Taylor & Francis LTD, 1975.

[11] S. Smale. Topology and mechanics I. *Inventiones math.*, 10:305–331, 1970.

[12] Mikael Spivak. *A comprehensive introduction differential geometry*, volume one. PUBLISH OR PERISH, INC., third edition, 1999.