Geometric Mechanics - Part I

Bob Rink*

January 13, 2009

Contents

*Vrije Universiteit Amsterdam, Faculty of Science, Department of Mathematics, De Boelelaan 1081a, 1081 HV Amsterdam, e-mail `brink@few.vu.nl`

Introduction

Classical mechanics has had an enormous influence on mathematics. The development of differential calculus by Newton for instance was mainly inspired by his desire to understand the motion of the planets. But there are also deep connections between classical mechanics and differential geometry, topology, dynamical systems theory and the calculus of variations, to name a few.

My part of the course could be called "Lagrangian mechanics". It starts with the derivation of the Euler-Lagrange equations from a variational princible. As an example we will review geodesic motion.

We will then focus on mechanics on Lie groups. It is here that the geometry really enters the mechanics. We will encounter the "Euler-Poincaré" symmetry reduction and the famous rigid body motion. We will moreover study an infinite-dimensional example that combines all these concepts: the Euler equations for an ideal incompressible fluid.

Finally, as an introduction to the remainder of this course, I will introduce the concept of a Hamiltonian system.

Bob Rink
Amsterdam, January 2009

1 Mechanical systems

The most important principle in classical mechanics is the property that a mechanical system can be given an arbitrary initial position and velocity, but that these then determine the behaviour of the system completely.

Let us see what this means. Let $q : I \to U, t \mapsto q(t)$ be a C^2 curve in an open subset $U \subset \mathbb{R}^n$, defined for t in an open interval $I \subset \mathbb{R}$. We assume that $q(t)$ describes the position or "configuration" of a mechanical system at time t. The velocities of the mechanical system are then given by the derivatives

$$\dot{q}_j(t) := \frac{dq_j(t)}{dt} \in \mathbb{R} \ , \ j = 1, \ldots, n \ .$$

The above main principle of classical mechanics then implies that the accelerations at time t,

$$\frac{d\dot{q}_j(t)}{dt} = \frac{d^2 q_j(t)}{dt^2} \ ,$$

are determined by the positions and velocities at time t, that is

$$\frac{d\dot{q}_j(t)}{dt} = a_j\left(t, q(t), \dot{q}(t)\right) \ ,$$

for certain functions $a_j : I \times U \times \mathbb{R}^n \to \mathbb{R}$. The exact form of the functions a_j depends on the physical properties of the mechanical system under consideration.

The ordinary differential equations

$$\begin{aligned} \frac{dq_j}{dt} &= \dot{q}_j \ , \\ \frac{d\dot{q}_j}{dt} &= a_j(t, q, \dot{q}) \ . \end{aligned} \tag{1.1}$$

are defined on the "phase space" $U \times \mathbb{R}^n$ and are called the "equations of motion" for the mechanical system under consideration. Under the mild condition that the a_j are locally Lipschitz continuous, the existence and uniqueness theory for ordinary differential equations indeed guarantees that $q(t)$ is determined by these equations of motion once the initial position $q(t_0)$ and velocity $\dot{q}(t_0)$ are given.

1.1 Two classical examples

A simple but famous example of a classical mechanical system was found by Galilei. He discovered experimentally that falling objects accelerate constantly towards the earth. In other words, Galilei discovered that falling objects describe a curve $q : t \mapsto q(t) \in \mathbb{R}^3$ that satisfies equations (1.1) with

$$a_1(t, q, \dot{q}) = a_2(t, q, \dot{q}) = 0 \ , \ a_3(t, q, \dot{q}) = -g \ . \tag{1.2}$$

The constant $g \in (0, \infty)$ is called the gravitational constant and it too can be determined experimentally: in Pisa it is approximately equal to $9,80\ m/s^2$. Integrating the equations $\frac{d\dot{q}_1}{dt} = \frac{d\dot{q}_2}{dt} = 0, \frac{d\dot{q}_3}{dt} = -g$ gives that $\dot{q}_1(t) = \dot{q}_1(0), \dot{q}_2(t) = \dot{q}_2(0), \dot{q}_3(t) = \dot{q}_3(0) - gt$. Integrating also $\frac{dq_i}{dt} = \dot{q}_i$, we then find that

$$q_1(t) = q_1(0) + \dot{q}_1(0)t, q_2(t) = q_2(0) + \dot{q}_2(0)t, q_3(t) = q_3(0) + \dot{q}_3(0)t - \frac{1}{2}gt^2 \ .$$

Thus, one can even explicitly solve Galiei's equations of motion.

A more complicated but equally famous mechanical system is the Kepler system that describes the motion of a planet around the sun, given by a curve $t \mapsto q(t) \in \mathbb{R}^2$. Based on observations by the astronomer Tycho Brahe, Kepler formulated the following principles for this motion, now known as *Kepler's laws*:

1. A planet moves on an ellipse in $\mathbb{R}^2$, with the sun in one of its focal points. Call this point $q_s \in \mathbb{R}^2$.

2. The area of the domain bounded by the line segment from q_s to $q(t_0)$, the orbit of the planet and the line segment from q_s to $q(t)$ is proportional to $t - t_0$.

3. The square of the period of a planetary orbit divided by the third power of the length of the long axis of its elliptic orbit is the same for every planet.

Newton proved that Kepler's laws are in fact equivalent to the equations of motion

$$m\frac{d^2q}{dt^2} = -\frac{mMG}{||q - q_s||^3}(q - q_s) \ , \tag{1.3}$$

in which m is the mass of the planet, M is the mass of the sun and G is a universal gravitational constant that is the same for every planet. Having at our disposal the techniques of modern calculus, the proof of Newton's theorem is quite elementary, but we will not present it here. In fact, it is well-known that the solutions to equation (1.3) describe a *conic section*, i.e. a circle, ellipse, parabola or hyperbola.

The right hand side of equation (1.3) is called the *force* acting on the planet, so equation (1.3) illustrates *Newton's first law* that says that the mass of a body times its acceleration is equal to the force acting on the body. Newton's *second law* asserts that the total force acting on a body is equal to the sum of the forces that are acting on it. So for instance, if $t \mapsto q^{(i)}(t) \in \mathbb{R}^2$ $(i = 1, \ldots, N)$ describe the positions of a collection of N planets moving in $\mathbb{R}^2$, each with their own mass m_i, then the equations of motion for these planets are given by

$$m_i\frac{d^2q^{(i)}}{dt^2} = \sum_{j \neq i} \phi_{ji}(q) \ , \tag{1.4}$$

where

$$\phi_{ji}(q) := -\frac{m_i m_j G}{||q^{(i)} - q^{(j)}||^3}(q^{(i)} - q^{(j)}) \tag{1.5}$$

is the force that planet j exerts on planet i. We also see here an illustration of Newton's *third law*, which says that the force that one body exerts on another body is equal to minus the force that this other body exerts on the first body. For the motion of the planets this is expressed by the fact that $\phi_{ji} = -\phi_{ij}$ as is clear from formula (1.5).

1.2 A constant of motion

A general mechanical system with n degrees of freedom is defined by Newton's system of second order ordinary differential equations

$$m_i \frac{d^2 q_i}{dt^2} = \phi_i(q) \ , \ i = 1, \ldots, n \ . \tag{1.6}$$

Here, $q \in U \subset \mathbb{R}^n$ is an element of an open subset of $\mathbb{R}^n$ and the continuous functions $\phi_i : U \to \mathbb{R}$ are the components of the "force" acting on q. Note that we assumed that the ϕ_i only depend on the positions q, just as in the examples of the previous paragraph.

Defining again the velocities $\dot{q}_i := \frac{dq_i}{dt}$, this system of n second order equations is equivalent to the system of $2n$ first order equations on $U \times \mathbb{R}^n$

$$\begin{aligned} \frac{dq_i}{dt} &= \dot{q}_i \\ m_i \frac{d\dot{q}_i}{dt} &= \phi_i(q) \ , \ i = 1, \ldots, n \ . \end{aligned} \tag{1.7}$$

Suppose now that for some reason there exists a C^1 function $V : U \to \mathbb{R}$ with the property that

$$\phi_i(q) = -\frac{\partial V(q)}{\partial q_i} \ .$$

If this is the case, then we call V the potential of the force ϕ and ϕ is called a conservative force. In general, the requirement that a force be conservative is very restrictive. For instance, if the ϕ_i are C^1, then V is C^2 and because $\frac{\partial^2 V}{\partial q_i \partial q_j} = \frac{\partial^2 V}{\partial q_j \partial q_i}$, the requirement that $\phi_i = -\frac{\partial V}{\partial q_i}$ for all i leads to the conclusion that

$$\frac{\partial \phi_i}{\partial q_j} = \frac{\partial \phi_j}{\partial q_i} \ , \ i, j = 1, \ldots, n \ . \tag{1.8}$$

It turns out the condition (1.8) is sufficient for the existence of an open subset U_{q_0} near each $q_0 \in U$ and a C^2 function $V_{q_0} : U_{q_0} \to \mathbb{R}$ such that $\phi_i = -\frac{\partial V_{q_0}}{\partial q_i}$ on U_{q_0}. The existence of such a function on the entire U can only be guaranteed under strong topological conditions on U, for instance that it be starshaped or, more generally, simply connected. In fact, the very question when a force is conservative was the motive for Poincaré to introduce the subject of topology.

The reason why conservative forces are so important lies in the fact that this property implies that the function of positions and velocities

$$E(q, \dot{q}) := \sum_{i=1}^{n} \frac{1}{2} m_i \dot{q}_i^2 + V(q)$$

is a so-called "constant of motion" for the system (1.7), because along a solution $(q(t), \dot{q}(t))$ of (1.7),

$$\frac{d}{dt}E\left(q(t), \dot{q}(t)\right) = \sum_{i=1}^{n}\left(\frac{\partial E(q,\dot{q})}{\partial q_i}\frac{dq_i(t)}{dt} + \frac{\partial E(q,\dot{q})}{\partial \dot{q}_i}\frac{d\dot{q}_i(t)}{dt}\right) =$$

$$\sum_{i=1}^{n}\left(\frac{\partial E(q,\dot{q})}{\partial q_i}\dot{q}_i + \frac{\partial E(q,\dot{q})}{\partial \dot{q}_i}\frac{\phi_i(q)}{m_i}\right) = \sum_{i=1}^{n}\left(-\phi_i(q)\dot{q}_i + m_i\dot{q}_i\frac{\phi_i(q)}{m_i}\right) = 0 \ .$$

In other words, the level sets of the function E are invariant under the flow of (1.7). At regular points, the level set of E is a submanifold of $U \times \mathbb{R}^n$ of dimension $2n-1$. The singular points of E correspond to the equilibrium points of (1.7).

Remark 1.1 (Energy) The function $V = V(q)$ is called the *potential energy* of a conservative mechanical system. The function

$$T(\dot{q}) := \sum_{i=1}^{n}\frac{1}{2}m_i\dot{q}_i^2 \tag{1.9}$$

is called *kinetic energy* and the function $E(q,\dot{q}) = T(\dot{q}) + V(q)$ is called the *total energy* of the mechanical system. The concept of conservation of total energy is of course quite fundamental in classical physics.

Remark 1.2 (Friction) Including in equation (1.6) a linear friction term leads to the equation

$$m_i\frac{d^2q_i}{dt^2} = \phi_i(q) - c\frac{dq_i}{dt} \ ,$$

where $c > 0$ is called the friction constant. For solutions to these equations one computes that

$$\frac{d}{dt}E\left(q(t), \frac{dq(t)}{dt}\right) = -c||\frac{dq(t)}{dt}||^2 \ ,$$

which means that E decreases along solutions of the differential equation, unless $q(t)$ is constant. We say that in this case, E is a *Lyapunov function.*

1.3 Exercises

Exercise 1.1 (Galilei's laws with friction) *If we include in Galilei's model for falling objects a friction force $\phi(q,\dot{q}) = -c\dot{q}$ that acts in the direction of minus the velocity $\dot{q}$ and is proportional to $||\dot{q}||$, then his equations become*

$$\frac{dq_1}{dt} = \dot{q}_1 \ , \frac{dq_2}{dt} = \dot{q}_2, \frac{dq_3}{dt} = \dot{q}_3, m\frac{d\dot{q}_1}{dt} = -c\dot{q}_1, m\frac{d\dot{q}_2}{dt} = -c\dot{q}_2, m\frac{d\dot{q}_3}{dt} = -mg - c\dot{q}_3 \ .$$

$c > 0$ is called the friction constant. Solve these equations and show that $\lim_{t\to\infty}(\dot{q}_1, \dot{q}_2, \dot{q}_3) = (0, 0, -\frac{mg}{c})$*. Hence we observe that in the presence of friction, a falling object does not accelerate without bound, but instead approaches a limiting speed.*

Exercise 1.2 (Charged particle) *The motion of a charged test particle in an electric and magnetic field is given by a curve $q(t) \in \mathbb{R}^3$ satisfying*

$$m\frac{d^2q}{dt^2} = e\left(E + \frac{dq}{dt} \times B\right) ,$$

where $E = E(q,t) \in \mathbb{R}^3$ and $B = B(q,t) \in \mathbb{R}^3$ are vector functions, called the electric and magnetic field respectively and $e \in \mathbb{R}$ is called the charge of the particle. Show that $\langle \frac{d^2q}{dt^2}, B\rangle = 0$ if $E = 0$. Under the assumption that $E = 0$ and $B = (b,0,0) \in \mathbb{R}^3$ is constant, solve the equations of motion.

Exercise 1.3 *Show that Galilei's equations (1.2) are conservative. Give T, V and E. Apart from E, can you point out other constants of motion?*

Exercise 1.4 *According to Newton's laws, the motion of a system of N planets is given by*

$$m_i\frac{d^2q^{(i)}}{dt^2} = \sum_{j\neq i} \phi_{ji}(q) ,$$

where ϕ_{ji} is given in equation (1.5). Let $M := \sum_{i=1}^N m_i$ be the total mass of the planets and define the center of mass as

$$z(t) := \frac{1}{M}\sum_{i=1}^N m_i q^{(i)}(t) .$$

Prove that $z(t)$ is a linear function of t. Show that the components of

$$\mu := \frac{dz}{dt} = \frac{1}{M}\sum_{i=1}^N m_i \dot{q}^{(i)}$$

are constants of motion.

Exercise 1.5 *If we define, for $i = 1,\ldots,N$ and $k = 1,2$, the total force acting on the k-th coordinate of planet i, by $\phi_k^{(i)}(q) = \sum_{j\neq i}(\phi_{ji}(q))_k$, then the equations of motion (1.4) can be written as*

$$m_i\frac{d^2q_k^{(i)}}{dt^2} = \phi_k^{(i)}(q) .$$

Show that

$$\phi_k^{(i)}(q) = -\frac{\partial V(q)}{\partial q_k^{(i)}} ,$$

with

$$V(q) := -G\sum_{i<j}\frac{m_i m_j}{||q^{(i)} - q^{(j)}||} .$$

This shows that the forces acting on the planets are conservative. Give the constant of motion.

2 Lagrangian mechanics

Lagrange showed that Newton's equations (1.6) are defined in a coordinate-invariant way. More precisely, he formulated Newton's equation in such a way that they behave well under coordinate transformations. Lagrange's construction will be the main topic of this section.

2.1 New position variables

One may wonder what happens to the equations of motion (1.6) for $q(t) \in U \subset \mathbb{R}^n$ if we make an arbitrary change of the position variables, that is if assume that $q(t) = \Phi(t, Q(t))$ for $Q(t) \in \tilde{U} \subset \mathbb{R}^m$. The reason we ask this question is that, once we know how an arbitrary Φ changes the equations of motion, we might be able to make a clever choice of Φ that changes the equations (1.6) for $q(t)$ into much simpler equations for $Q(t)$.

If Φ is C^2, then a twice differentiable curve $t \mapsto Q(t)$ in $\tilde{U} \subset \mathbb{R}^m$ defined on some open time-interval $I \subset \mathbb{R}$ is transformed by Φ into a twice differentiable curve

$$q(t) := \Phi(t, Q(t))$$

in $U \subset \mathbb{R}^n$. Differentiation of $q(t)$ gives that

$$\frac{dq_i(t)}{dt} = \frac{\partial \Phi_i(t, Q(t))}{\partial t} + \sum_{j=1}^{m} \frac{\partial \Phi_i(t, Q(t))}{\partial Q_j} \frac{dQ_j(t)}{dt} \ ,$$

i.e. $\frac{dq}{dt} = \frac{\partial \Phi}{\partial t} + \frac{\partial \Phi}{\partial Q}\frac{dQ}{dt}$, which is a nice *affine* transformation formula. Differentiating this once more we nevertheless find that even if Φ does not depend explicitly on t, the second order derivative $\frac{d^2q}{dt^2}$ in general is not simply the image of $\frac{d^2Q}{dt^2}$ under the linear map $\frac{\partial \Phi}{\partial Q}$.

Lagrange realised that one can take another approach, which might at first seem a little artificial. The first thing he remarked is that

$$m_i \frac{d^2 q_i(t)}{dt^2} = \frac{d}{dt}\left(\left.\frac{\partial T(\dot{q})}{\partial \dot{q}_i}\right|_{\dot{q}=\frac{dq(t)}{dt}} \right) \ , \tag{2.1}$$

where T is the kinetic energy defined in (1.9). The next remark is that one can express the kinetic energy $T\left(\frac{dq(t)}{dt}\right)$ explicitly as a function of t, $Q(t)$ and $\frac{dQ(t)}{dt}$, by defining the function $\tilde{T} : I \times \tilde{U} \times \mathbb{R}^m \to \mathbb{R}$ by

$$\tilde{T}(t, Q, \dot{Q}) = T(\dot{q}) \ ,$$

in which

$$\dot{q} = \frac{\partial \Phi(t, Q)}{\partial t} + \frac{\partial \Phi(t, Q)}{\partial Q} \cdot \dot{Q} \ .$$

In other words, $\tilde{T}$ is defined by $\tilde{T}(t, Q, \dot{Q}) = T\left(\frac{\partial \Phi(t,Q)}{\partial t} + \frac{\partial \Phi(t,Q)}{\partial Q} \cdot \dot{Q}\right)$. This definition is such that $T\left(\frac{dq(t)}{dt}\right) = \tilde{T}\left(t, Q(t), \frac{dQ(t)}{dt}\right)$ if $q(t) = \Phi(t, Q(t))$. Note that the transformed kinetic energy $\tilde{T}(t,Q,\dot{Q})$ in general will depend explicitly on the time t and the new position variable Q. Now Lagrange's discovery was that not $\frac{d}{dt}\left(\frac{\partial\tilde{T}(t,Q,\dot{Q})}{\partial\dot{Q}_j}\Big|_{Q=Q(t),\dot{Q}=\frac{dQ(t)}{dt}}\right)$, but the quantity $\frac{d}{dt}\left(\frac{\partial\tilde{T}(t,Q,\dot{Q})}{\partial\dot{Q}_j}\Big|_{Q=Q(t),\dot{Q}=\frac{dQ(t)}{dt}}\right) - \frac{\partial\tilde{T}(t,Q,\dot{Q})}{\partial Q_j}\Big|_{Q=Q(t),\dot{Q}=\frac{dQ(t)}{dt}}$ depends linearly on $\frac{d}{dt}\left(\frac{\partial T(\dot{q})}{\partial\dot{q}_i}\Big|_{\dot{q}=\frac{dq(t)}{dt}}\right)$. In fact, we have the following quite general result:

Theorem 2.1 *Let $\Phi : I\times\tilde{U}\to U$ be a C^2 map, $t\mapsto Q(t)$ a C^2 curve in $\tilde{U}$ and $q(t) = \Phi(t,Q(t))$ the corresponding curve in U. Let $L : I\times U\times\mathbb{R}^n\to\mathbb{R}$ be a C^2 function of $(t,q,\dot{q})$ and let $\tilde{L} : I\times\tilde{U}\times\mathbb{R}^m\to\mathbb{R}$ be the corresponding function of $(t,Q,\dot{Q})$ defined by*

$$\tilde{L}(t,Q,\dot{Q}) = L\left(t,\Phi(t,Q),\frac{\partial\Phi(t,Q)}{\partial t}+\frac{\partial\Phi(t,Q)}{\partial Q}\cdot\dot{Q}\right)\ . \tag{2.2}$$

Furthermore, define the continuous functions $[L]^q_i$ on I by

$$[L]^q_i(t) := \frac{d}{dt}\left(\frac{\partial L(t,q,\dot{q})}{\partial\dot{q}_i}\Big|_{\dot{q}=\frac{dq(t)}{dt}}\right) - \frac{\partial L(t,q,\dot{q})}{\partial q_i}\Big|_{\dot{q}=\frac{dq(t)}{dt}}\ ,\ i=1,\dots,n\ , \tag{2.3}$$

and similarly $[\tilde{L}]^Q_j$:

$$[\tilde{L}]^Q_j(t) := \frac{d}{dt}\left(\frac{\partial\tilde{L}(t,Q,\dot{Q})}{\partial\dot{Q}_j}\Big|_{\dot{Q}=\frac{dQ(t)}{dt}}\right) - \frac{\partial\tilde{L}(t,Q,\dot{Q})}{\partial Q_j}\Big|_{\dot{Q}=\frac{dQ(t)}{dt}}\ ,\ j=1,\dots,m\ . \tag{2.4}$$

Then we have the following transformation formula:

$$[\tilde{L}]^Q_j(t) = \sum_{i=1}^n [L]^q_i(t)\cdot\frac{\partial\Phi_i(t,Q)}{\partial Q_j}\Big|_{Q=Q(t)}\ ,\ j=1,\dots,m\ . \tag{2.5}$$

One can prove this theorem by a very long direct computation, using the relations $q = \Phi(t,Q)$ and $\frac{dQ}{dt} = \frac{\partial\Phi(t,Q)}{\partial t}+\frac{\partial\Phi(t,Q)}{\partial Q}\frac{dq}{dt}$ and differentiating the identity (2.2) with respect to Q_j and $\dot{Q}_j$. But one can also prove Theorem 2.1 in a surprisingly different way, as we shall see in the next section.

The conclusion of Theorem 2.1 is that $[\tilde{L}]^Q$ depends linearly on $[L]^q$, namely $[\tilde{L}]^Q = [L]^q\cdot\frac{\partial\Phi}{\partial Q}$, viewing $[L]^q(t)$ and $[\tilde{L}]^Q(t)$ as row vectors. Note that this is *not* the same as the transformation formula $\frac{dq}{dt} = \frac{\partial\Phi}{\partial t}+\frac{\partial\Phi}{\partial Q}\cdot\frac{dQ}{dt}$ for the velocity vectors, even if Φ is independent of t. The classical terminology is that $[L]^q$ transforms *covariantly* and $\frac{dq}{dt}$ transforms *contravariantly*.

2.2 A variational proof of Theorem 2.1

This section contains a "variational" proof of Theorem 2.1. It will become clear what this means. I myself think that this proof is quite surprising.

We start with letting $q : I \to U \subset \mathbb{R}^n$ be a C^2 curve in U and $L : I \times U \times \mathbb{R}^n \to \mathbb{R}$ a C^2 Lagrangian function. Then we can integrate L over compact pieces of the curve $t \mapsto q(t)$. Hence we fix some $a, b \in I, a < b$ and define the *action* of L along q restricted to $[a, b]$ as

$$A(q) := \int_a^b L\left(t, q(t), \frac{dq(t)}{dt}\right) dt \ .$$

The next step is to consider small perturbations of the curve $t \mapsto q(t)$, depending on an auxiliary parameter ε. That is we consider C^2 maps

$$\tilde{q} : I \times (-\varepsilon_0, \varepsilon_0) \to U \ ,$$

with the property that $\tilde{q}(t, 0) = q(t)$. For an $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ close to 0, the curve $q_\varepsilon : t \mapsto \tilde{q}(t, \varepsilon)$ lies "close" to the curve $t \mapsto q(t)$, and hence such a map $\tilde{q}$ is called a *variation* of the curve $t \mapsto q(t)$. Now the action can be viewed as a function $\varepsilon \mapsto A(q_\varepsilon)$ on $(-\varepsilon_0, \varepsilon_0)$. Due to the theorem for interchanging differentiation and integration, this function is itself differentiable and differentiation gives:

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} A(q_\varepsilon) = \int_a^b \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} L\left(t, q_\varepsilon(t), \frac{dq_\varepsilon(t)}{dt}\right) dt =$$

$$\int_a^b \sum_{i=1}^n \left(\left.\frac{\partial L(t, q, \dot{q})}{\partial q_i}\right|_{\substack{q = q(t) \\ \dot{q} = \frac{dq(t)}{dt}}} \left.\frac{\partial \tilde{q}_i(t, \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} + \left.\frac{\partial L(t, q, \dot{q})}{\partial \dot{q}_i}\right|_{\substack{q = q(t) \\ \dot{q} = \frac{dq(t)}{dt}}} \left.\frac{\partial^2 \tilde{q}_i(t, \varepsilon)}{\partial t \partial \varepsilon}\right|_{\varepsilon=0} \right) dt \ .$$

Partial integration with respect to t of the second term in this expression now gives that this is equal to

$$-\int_a^b \sum_{i=1}^n [L]_i^q(t) \left.\frac{\partial \tilde{q}_i(t, \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} dt + \text{ "boundary terms"} \ .$$

If we consider only variations $\tilde{q}$ of q with fixed endpoints, that is $\tilde{q}(a, \varepsilon) = \tilde{q}(a, 0) = q(a)$ and $\tilde{q}(b, \varepsilon) = \tilde{q}(b, 0) = q(b)$, then $\frac{\partial \tilde{q}(a,\varepsilon)}{\partial \varepsilon} = \frac{\partial \tilde{q}(b,\varepsilon)}{\partial \varepsilon} = 0$, whence the boundary terms disappear. From now on we will assume that our variations have fixed endpoints.

If $q(t) = \Phi(t, Q(t))$ with $\Phi : I \times \tilde{U} \to U$, and $\tilde{Q} : I \times (-\varepsilon_0, \varepsilon_0) \to \tilde{U}$ is a variation of $Q(t)$ with fixed endpoints, consisting of curves $Q_\varepsilon(t) := \tilde{Q}(t, \varepsilon)$, then $\tilde{q}(t, \varepsilon) := \Phi(t, \tilde{Q}(t, \varepsilon))$ is a variation of $t \mapsto q(t)$ with fixed endpoints. The definition (2.2) of $\tilde{L}$ implies that $\tilde{L}\left(t, Q_\varepsilon(t), \frac{dQ_\varepsilon(t)}{dt}\right) = L\left(t, q_\varepsilon(t), \frac{dq_\varepsilon(t)}{dt}\right)$, so that

$$A(q_\varepsilon) = \tilde{A}(Q_\varepsilon) := \int_a^b \tilde{L}\left(t, Q_\varepsilon(t), \frac{dQ_\varepsilon(t)}{dt}\right) dt \ .$$

Thus, differentiation of the identity $A(q_\varepsilon) = \tilde{A}(Q_\varepsilon)$ with respect to ε at $\varepsilon = 0$ gives that

$$\int_a^b \sum_{i=1}^n [L]_i^q(t) \left.\frac{\partial \tilde{q}_i(t,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} dt = \int_a^b \sum_{j=1}^m [\tilde{L}]_j^Q(t) \left.\frac{\partial \tilde{Q}_j(t,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} dt \ . \tag{2.6}$$

On the other hand, differentiation of $\tilde{q}_i(t,\varepsilon) = \Phi_i(t, \tilde{Q}(t,\varepsilon))$ gives that

$$\left.\frac{\partial \tilde{q}_i(t,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = \sum_{j=1}^m \left.\frac{\partial \Phi_i(t,Q)}{\partial Q_j}\right|_{Q=Q(t)} \left.\frac{\partial \tilde{Q}_j(t,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} ,$$

and we conclude that

$$\int_a^b \sum_{j=1}^m \left([\tilde{L}]_j^Q(t) - \sum_{i=1}^n [L]_i^q(t) \left.\frac{\partial \Phi_i(t,Q)}{\partial Q_j}\right|_{Q=Q(t)} \right) \left.\frac{\partial \tilde{Q}_j(t,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} dt = 0 \ , \tag{2.7}$$

for every C^2 variation $\tilde{Q}$ of Q.

Finally, suppose that $[\tilde{L}]_j^Q(t) \neq \sum_{i=1}^n [L]_i^q(t) \left.\frac{\partial \Phi_i(t,Q)}{\partial Q_j}\right|_{Q=Q(t)}$ for some $1 \leq j \leq m$ and at some $t = t^*$ with $a < t^* < b$. Then because of continuity, inequality holds in an interval $[t^* - \delta, t^* + \delta]$. If we now choose a variation $\tilde{Q}$ of Q with $\tilde{Q}_k(t,\varepsilon) = Q_k(t)$ for $k \neq j$ and $\tilde{Q}_j(t,\varepsilon) = Q_j(t) + \varepsilon\chi(t)$, with χ some nonzero C^2 function of fixed sign and with compact support in $[t^* - \delta, t^* + \delta]$, then for this variation, formula (2.7) is not true: a contradiction. This proves Theorem 2.1.

2.3 Euler-Lagrange equations

For a given C^2 function $L : I \times U \times \mathbb{R}^n \to \mathbb{R}$, the equations of motion

$$[L]^q = 0$$

for the curve $t \mapsto q(t)$ in U are called the *Euler-Lagrange equations* for L and L is called the *Lagrangian function* for the equations $[L]^q = 0$.

A particular consequence of Theorem 2.1 is that if $t \mapsto q(t)$ solves the Euler-Lagrange equations for L and $q(t) = \Phi(t, Q(t))$, then $t \mapsto Q(t)$ automatically solves the Euler-Lagrange equations for $\tilde{L}$. If for all $t \in I$, $\Phi(t, \cdot)$ is a diffeomorphism, that is if $\frac{\partial \Phi}{\partial Q}$ is invertible, then the reverse statement is also true. This expresses that Euler-Lagrange equations for $q(t)$ are defined coordinate-invariantly, which is hardly surprising because they are equivalent to a *variational principle.*

More precisely, the proof of Theorem 2.1 shows that $[L]^q = 0$ if and only if

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} A(\tilde{q}(\cdot,\varepsilon)) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_a^b L\left(t, \tilde{q}(t,\varepsilon), \frac{\partial \tilde{q}(t,\varepsilon)}{\partial t}\right) dt = 0$$

for all $a, b \in I$ and all C^2 maps $\tilde{q} : I \times (-\varepsilon_0, \varepsilon_0) \to U$ with the property that

1. $\tilde{q}$ is a variation of q, i.e. $\tilde{q}(t,0) = q(t)$.

2. $\tilde{q}$ has fixed endpoints, i.e. $\tilde{q}(a,\varepsilon) = q(a)$ and $\tilde{q}(b,\varepsilon) = q(b)$.

In short, we say that *the action A is stationary at the curve $t \mapsto q(t)$ for variations with fixed endpoints*. Some authors express this by writing

$$\delta A(q) = \delta \int_a^b L\left(t, q(t), \frac{dq(t)}{dt}\right) dt = 0 \ ,$$

thus expressing that the "derivative" of A at the curve $t \mapsto q(t)$ is zero. We will not use this notation.

The statement that $t \mapsto q(t)$ is a solution of the equations $[L]^q = 0$ if and only if it is stationary for $A(q) = \int_a^b L(t, q(t), \frac{dq(t)}{dt})dt$ for variations with fixed endpoints is known as *Hamilton's principle.*

Let us study the Euler-Lagrange equations for L in some more detail now. Written out explicitly, they read $\frac{dq_i}{dt} = \dot{q}_i$ and

$$\frac{\partial^2 L(t,q,\dot{q})}{\partial t \partial \dot{q}_i} + \sum_{j=1}^n \left(\frac{\partial^2 L(t,q,\dot{q})}{\partial q_j \partial \dot{q}_i} \dot{q}_j + \frac{\partial^2 L(t,q,\dot{q})}{\partial \dot{q}_j \partial \dot{q}_i} \frac{d\dot{q}_j}{dt} \right) - \frac{\partial L(t,q,\dot{q})}{\partial q_i} = 0 \ , \ i = 1, \ldots, n \ . \tag{2.8}$$

If the second order derivative matrix $\left.\frac{\partial^2 L(t,q,\dot{q})}{\partial \dot{q}^2}\right|_{\substack{q \,=\, q(t) \\ \dot{q} \,=\, \frac{dq(t)}{dt}}}$ is invertible, then we call L a *non-degenerate Lagrangian* at $(t, q, \dot{q})$. Nondegeneracy implies that near $(t, q, \dot{q})$ we can rewrite the Euler-Lagrange equations explicitly as a system $\frac{dq_i}{dt} = \dot{q}_i$, $\frac{d\dot{q}_i}{dt} = \tilde{\phi}_i(t, q, \dot{q})$, and in particular we then have local existence and uniqueness of the solutions to the Euler-Lagrange equations. The requirement that $\frac{\partial^2 L}{\partial \dot{q}^2}$ is invertible is called the *Legendre condition.*

The final remark in this section is that if, for a C^2 Lagrangian function $L = L(t, q, \dot{q})$, we define the function $h = \frac{\partial L}{\partial \dot{q}} \cdot \dot{q} - L$, that is

$$h(t,q,\dot{q}) := \sum_{i=1}^n \frac{\partial L(t,q,\dot{q})}{\partial \dot{q}_i} \dot{q}_i - L(t,q,\dot{q}) \ , \tag{2.9}$$

then it is easy to compute that

$$\frac{d}{dt} h\left(t, q(t), \frac{dq(t)}{dt}\right) = - \left.\frac{\partial L(t,q,\dot{q})}{\partial t}\right|_{\substack{q \,=\, q(t) \\ \dot{q} \,=\, \frac{dq(t)}{dt}}} + \sum_{i=1}^n [L]_i^q(t) \frac{dq_i(t)}{dt} \ . \tag{2.10}$$

Corollary 2.2 *If $t \mapsto q(t)$ is a solution to the Euler-Lagrange equations $[L]^q = 0$ and L does not explicitly depend on time, then $h := \frac{\partial L}{\partial \dot{q}} \cdot \dot{q} - L$ is constant along $t \mapsto q(t)$.*

2.4 Natural mechanical systems

Let us now return to Newton's equations of motion in conservative form,

$$m_i \frac{d^2 q_i}{dt} = -\frac{\partial V(q)}{\partial q_i} \ , \tag{2.11}$$

for which we defined the kinetic energy $T(\dot{q}) = \sum_{i=1}^{n} \frac{1}{2} m_i \dot{q}_i^2$, the potential energy $V(q)$ and the total energy $E = T + V$. The interesting remark is that if we also define the *Lagrangian* function

$$L : U \times \mathbb{R}^n \ , \ L(q, \dot{q}) = T(\dot{q}) - V(q) \ ,$$

of kinetic energy *minus* potential energy, then

$$[L]_i^q(t) = m_i \frac{d^2 q_i(t)}{dt^2} + \left. \frac{\partial V(q)}{\partial q_i} \right|_{q=q(t)} \ .$$

In other words, $t \mapsto q(t)$ solves Newton's equations of motion (2.11) if and only if $[T - V]^q = 0$.

Moreover, if $L(q, \dot{q}) = T(\dot{q}) - V(q)$, with $T(\dot{q}) = \sum_{i=1}^{n} \frac{1}{2} m_i \dot{q}_i^2$, then $\sum_{i=1}^{n} \frac{\partial L(q,\dot{q})}{\partial \dot{q}_i} \dot{q}_i = \sum_{i=1}^{n} \frac{\partial T(\dot{q})}{\partial \dot{q}_i} \dot{q}_i = 2T(\dot{q})$, so that

$$h(q, \dot{q}) = T(\dot{q}) + V(q) = E(q, \dot{q}) \ ,$$

and we find that h is equal to our old constant of motion, the total energy E.

If $T(\dot{q})$ is the kinetic energy function as above and $q = \Phi(Q)$, $\dot{q} = \frac{\partial \Phi(Q)}{\partial Q} \cdot \dot{Q}$, then the transformed kinetic energy $S(Q, \dot{Q}) = T(\dot{q})$ is of the form

$$S(Q, \dot{Q}) = \frac{1}{2} \sum_{i,j=1}^{n} \beta_{ij}(Q) \dot{Q}_i \dot{Q}_j$$

for certain functions $Q \mapsto \beta_{ij}(Q)$ that are symmetric in the sense that $\beta_{ij} = \beta_{ji}$ for all $1 \leq i, j \leq n$.

Some terminology: Mechanical systems with a Lagrangian of the form $L(q, \dot{q}) = S(q, \dot{q}) - V(q)$, with S quadratic in $\dot{q}$, are sometimes called *natural* mechanical systems, because they can arise from Newton's equations possibly after a coordinate change. It should be clear now that the definition of a natural mechanical system does not depend on the choice of coordinates. As before, the function S is called the kinetic energy or free energy of the natural system and V the potential energy. By the above remarks, the total energy $E := S + V$ of a natural mechanical system is conserved.

2.5 Lagrangian equations for continua

Using Hamilton's principle as a physical postulate, we can derive equations of motion for the evolution of continuous media such as gases, fluids or elastic solids.

A continuum is modeled by a map $\psi : X \to \mathbb{R}^n$, where $X \subset \mathbb{R}^m$ is an open subset of $\mathbb{R}^m$. X is called the reference configuration of the continuum, and for $x \in X$, $\psi(x)$ is the location in $\mathbb{R}^n$ of the element of the continuum with label x. A *motion* of the continuum is a smooth map $u : I \times X \to \mathbb{R}^n$, where $I \subset \mathbb{R}$ is an open time-interval. Then the map $u(t, \cdot) : X \to \mathbb{R}^n$ describes the configuration of the continuum at time t, whereas the curve $t \mapsto u(t, x), I \to \mathbb{R}^n$ describes the motion of the element of the continuum with label x.

If $\rho : X \to \mathbb{R}$ is a mass density function, then the kinetic energy of the continuum is obtained by integrating the kinetic energy density $\frac{1}{2}\rho(x)||\dot{u}(x)||^2$ over X, where we denoted $\dot{u}(x) = \frac{\partial u(t,x)}{\partial t}$:

$$T\left(\dot{u}\right) = \int_X \frac{1}{2}\rho(x)||\dot{u}(x)||^2 \; d_m x \ .$$

The potential energy of the continuum at time t depends on the function $x \mapsto u(t, x)$ and may depend on the (x-)derivatives of this function. It is usually obtained by integrating a potential energy density function. For instance if we assume that it costs energy to stretch the continuum but not to bend it, then we are saying that the potential energy density depends only on the values of x, $u(t, x)$ and the matrix of first derivatives $Du(t, x) := \left(\frac{\partial u_i(t,x)}{\partial x_j}\right)_{i,j}$, and not on higher order derivatives. That is

$$V(u(t, \cdot)) = \int_X W(x, u(t, x), Du(t, x)) d_m x \ ,$$

in which $W : X \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \to \mathbb{R}$ is a smooth function. Let us denote the arguments of W by (x, u, B), where $x \in X, u \in \mathbb{R}^n, B \in \mathbb{R}^{n\times m}$.

Now Hamilton's principle means that $u : I \times X \to \mathbb{R}^n$ satisfies

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_a^b \int_X \frac{1}{2}\rho(x) \left|\left|\frac{\partial \tilde{u}(t, x, \varepsilon)}{\partial t}\right|\right|^2 - W(x, \tilde{u}(t, x, \varepsilon), D\tilde{u}(t, x, \varepsilon)) \; d_m x \; dt \; = 0$$

for all smooth variations $\tilde{u} : I \times X \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^n$ of u with fixed (temporal) endpoints. Let us choose the variations of the form $\tilde{u}(t, x, \varepsilon) = u(t, x) + \varepsilon\phi(t, x)$, where ϕ is smooth and has a compact support that is contained in $I \times X$. Bringing the derivative inside the integral we then obtain

$$\begin{aligned} \int_a^b \int_X \sum_i \rho(x) \frac{\partial u_i(t, x)}{\partial t} \frac{\partial \phi_i(t, x)}{\partial t} - \sum_i \left. \frac{\partial W(t, u, B)}{\partial u_i} \right|_{\substack{u \,=\, u(t,x) \\ B \,=\, Du(t,x)}} \cdot \phi_i(t, x) \\ - \sum_{i,j} \left. \frac{\partial W(t, u, B)}{\partial B_{ij}} \right|_{\substack{u \,=\, u(t,x) \\ B \,=\, Du(t,x)}} \cdot \frac{\partial \phi_i(t, x)}{\partial x_j} \; d_m x \; dt = 0 \ . \end{aligned} \tag{2.12}$$

If u is smooth enough, partial integration gives that this equals

$$\sum_i \int_a^b \int_X \phi_i(t, x) \left(-\rho(x) \frac{\partial^2 u_i(t, x)}{\partial t^2} - \left. \frac{\partial W(t, u, B)}{\partial u_i} \right|_{\substack{u=u(t,x) \\ B=Du(t,x)}} + \sum_j \frac{\partial}{\partial x_j} \left. \frac{\partial W(t, u, B)}{\partial B_{ij}} \right|_{\substack{u=u(t,x) \\ B=Du(t,x)}} \right) d_m x dt.$$

Thus we derived from Hamilton's principle the partial differential equations

$$\rho(x)\frac{\partial^2 u_i}{\partial t^2} + \frac{\partial W}{\partial u_i} - \sum_{j=1}^{m} \frac{\partial}{\partial x_j}\frac{\partial W}{\partial B_{ij}} = 0 \ , \ i = 1, \ldots, n \ . \tag{2.13}$$

Equations (2.13) form a system of n partial differential equations for the n functions u_i, depending explicitly on the second order derivatives of the u_i with respect to t and the first and second order derivatives of the u_i with respect to the x_j. If W had been a function of the second order derivatives D^2u as well, then the resulting partial differential equation for u would have included fourth-order x-derivatives of u, etc. The reader may want to try and derive these equations.

2.6 Exercises

Exercise 2.1 *Let $U \subset \mathbb{R}^n$ and $\tilde{U} \subset \mathbb{R}^m$ be open subsets and let $T : U \times \mathbb{R}^n \to \mathbb{R}$ and $V : U \to \mathbb{R}$ be C^2 functions, with $T(q, \dot{q}) = \sum_{i,j=1}^n \frac{1}{2}\beta_{ij}(q)\dot{q}_i\dot{q}_j$ with $\beta_{ij}(q) = \beta_{ji}(q)$ and let $L(q, \dot{q}) = T(\dot{q}) - V(q)$. Let $\Phi : \tilde{U} \to U$ be a C^2 map and let $\tilde{L}$ be defined by $\tilde{L}(Q, \dot{Q}) = L(\Phi(Q), \frac{\partial \Phi(Q)}{\partial Q} \cdot \dot{Q})$. Show that*

$$\tilde{L}(Q, \dot{Q}) = \frac{1}{2}\sum_{i,j=1}^{m} \frac{1}{2}\tilde{\beta}_{ij}(Q)\dot{Q}_i\dot{Q}_j - V(\Phi(Q)) \ ,$$

in which

$$\tilde{\beta}_{ij}(Q) = \sum_{k,l=1}^{n} \beta_{kl}(Q)\frac{\partial \Phi_k(Q)}{\partial Q_i}\frac{\partial \Phi_l(Q)}{\partial Q_j} \ .$$

Prove that $\tilde{\beta}_{ij}(Q) = \tilde{\beta}_{ji}(Q)$.

Exercise 2.2 (Rotating coordinates) *Let $U = \tilde{U} = \mathbb{R}^2$ and let $\Phi : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ be a one-parameter family of rotations: $q = \Phi(t, Q) = e^{t\sigma J} \cdot Q$, where σ is a real constant and*

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ .$$

- *If $T(\dot{q}) = \frac{1}{2}m\langle \dot{q}, \dot{q}\rangle$ is the kinetic energy of a free particle, compute the transformed kinetic energy $\tilde{T}(t, Q, \dot{Q})$. Show that it does not depend on t explicitly so that $\tilde{T} = \tilde{T}(Q, \dot{Q})$.*
- *Show that*

$$[\tilde{T}]^Q(t) = m\left(\frac{d^2Q(t)}{dt^2} + 2\sigma J\frac{dQ(t)}{dt} - \sigma^2 Q(t)\right) \ .$$

- *Let $V = V(q)$ be a rotation-symmetric potential energy function. Show that $\tilde{V}$ is independent of t and write down the Euler-Lagrange equations for $\tilde{L}(Q, \dot{Q}) = \tilde{T}(Q, \dot{Q}) - \tilde{V}(Q)$.*

Exercise 2.3 *We model a one-dimensional string by letting the reference and target configuration be one-dimensional:* $X = J \subset \mathbb{R}$ *is an open interval and for a smooth* $u : I \times J \to \mathbb{R}$, $u(t, \cdot)$ *describes the configuration of the string at time* t. *For a smooth mass density* $\rho : J \to \mathbb{R}$, *and velocity function* $\dot{u} : J \to \mathbb{R}$, *define the kinetic energy*

$$T(\dot{u}) = \int_J \frac{1}{2}\rho(x)\dot{u}(x)^2 dx \ .$$

Denoting $u^{(j)}(x) := \frac{\partial^j u(x)}{\partial x^j}$, *assume that the potential energy is given by*

$$V(u) = \int_J W(x, u(x), u^{(1)}(x), \ldots, u^{(k)}(x))dx \ ,$$

where W *is a smooth potential energy density function* $W : \mathbb{R}^{k+2} \to \mathbb{R}$. *Use Hamilton's principle to derive for* u *the partial differential equation*

$$\rho(x)\frac{\partial^2 u(t,x)}{\partial t^2} + \sum_{j=0}^{k}(-1)^j \frac{\partial^j}{\partial x^j} \left.\frac{\partial W(x, u, \ldots, u^{(k)})}{\partial u^{(j)}}\right|_{u=u(t,x),\ldots,u^{(k)}=u^{(k)}(t,x)} = 0 \ .$$

If $W = W(u^{(1)})$ *depends on the amount of "stretching" of the string only, we say that the string is purely elastic. For an elastic string derive the equation*

$$\rho(x)\frac{\partial^2 u(t,x)}{\partial t^2} = \frac{\partial^2 u(t,x)}{\partial x^2} W'\left(\frac{\partial u(t,x)}{\partial x}\right) \ .$$

Exercise 2.4 (Minimal surfaces) *Let* $X \subset \mathbb{R}^2$ *be a bounded open subset with a* C^2 *boundary* ∂X. *Let* $h : \overline{X} \to \mathbb{R}$ *be a continuous function that is* C^2 *on* X. *Then the graph of* h, $\mathrm{graph}(h) := \{(x, h(x)) \mid x \in X\}$ *is a smooth surface with a finite area that equals*

$$S(h) := \int_X || \begin{pmatrix} 1 \\ 0 \\ \frac{\partial h(x)}{\partial x_1} \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \frac{\partial h(x)}{\partial x_2} \end{pmatrix} || \, d_2x = \int_X \sqrt{1 + \left(\frac{\partial h(x)}{\partial x_1}\right)^2 + \left(\frac{\partial h(x)}{\partial x_2}\right)^2} \, d_2x \ .$$

We say that the graph of h *is a minimal surface over* X *if* $S(g) \geq S(h)$ *for all continuous functions* $g : \overline{X} \to \mathbb{R}$ *that are* C^2 *on* X *and equal to* h *on* ∂X. *In particular this implies that* $S(h + \varepsilon\phi) \geq S(h)$ *if* ϕ *is some* C^2 *function on* X *with compact support contained in* X. *Prove that this implies that on* X, h *satisfies the partial differential equation*

$$\left(1 + h_{x_1}^2 + h_{x_2}^2\right)(h_{x_1x_1} + h_{x_2x_2}) = h_{x_1}^2 h_{x_1x_1} + 2h_{x_1}h_{x_2}h_{x_1x_2} + h_{x_2}^2 h_{x_2x_2} \ .$$

Here we used the shorthand notation $h_\alpha := \frac{\partial h}{\partial \alpha}$. *This equation is called the* minimal surface equation.

3 The geodesic flow

In this chapter we will start making use of concepts from differential geometry. Where previously the configuration space of a mechanical system was an open subset U of $\mathbb{R}^n$, we now also allow the configuration space to be some n-dimensional differentiable manifold, denoted Q.

Recall that a differentiable manifold is a topological space that locally looks like an open subset of $\mathbb{R}^n$. As such, the open set $U \subset \mathbb{R}^n$ can be thought of as a coordinate patch for the manifold Q. Similarly, the space $U \times \mathbb{R}^n$ of positions and velocities is replaced by the tangent bundle TQ of Q. The reader is supposed to be sufficiently familiar with these concepts.

3.1 Riemannian manifolds

Riemannian manifolds were -no surprise- introduced by Riemann. His discovery was that by defining on every tangent space of a manifold an inner product, one can study concepts like curvature and shortest paths. In this chapter we will focus on the latter, the so-called geodesics.

Let us start with the notion of a *pseudo-inner product*:

Definition 3.1 *Let V be a finite-dimensional linear space. A pseudo-inner product on V is a mapping*

$$\beta : V \times V \to \mathbb{R}$$

with the following properties:

1. *Symmetry: $\beta(v, w) = \beta(w, v)$ for all $v, w \in V$.*
2. *Bilinearity: $\beta(v_1 + sv_2, w) = \beta(v_1, w) + s\beta(v_2, w)$ for all $v_1, v_2, w \in V$ and $s \in \mathbb{R}$.*
3. *Nondegeneracy: if $\beta(v, w) = 0$ for all $w \in V$, then $v = 0$.*

If in addition, we require

4. *Positivity: $\beta(v, v) \geq 0$,*

then β is called an inner product.

With this definition, we can define

Definition 3.2 *A C^k (pseudo-)Riemannian manifold is a C^k manifold Q for which a (pseudo-)inner product $\beta_q : T_qQ \times T_qQ \to \mathbb{R}$ is defined on each tangent space T_qQ, in such a way that this inner product depends in a C^k way on $q \in Q$.*

The family β_q of (pseudo-)inner products on T_qQ is called a *(pseudo-)Riemannian metric* on Q.

In induced local coordinate $U \times \mathbb{R}^n$ for the tangent bundle TQ, any (pseudo-)inner product takes the form

$$(\dot{q}^1, \dot{q}^2) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \sum_{i,j=1}^{n} \beta_{ij}(q)\dot{q}_i^1 \dot{q}_j^2 \ .$$

The matrix $\beta(q)$ with coefficients $\beta_{ij}(q)$ is symmetric, nondegenerate and, if we require positivity, positive definite. Moreover, it depends explicity on the local coordinates. The requirement that $\beta(q)$ depends in a C^k way on $q \in Q$ in Definition 3.2, just means that the functions $q \mapsto \beta_{ij}(q)$ on U are C^k.

Remark 3.3 (Quadratic forms) One can observe that a pseudo-inner product $\beta : V \times V \to \mathbb{R}$ on a linear space V defines a nondegenerate quadratic form $S_\beta : V \to \mathbb{R}$ by setting

$$S_\beta(v) := \frac{1}{2}\beta(v, v) \ .$$

On the other hand, given a nondegenerate quadratic form S on V, the unique pseudo-inner product β on V for which $S_\beta = S$, is given by

$$\beta_S(v, w) := \frac{1}{2}\left(S(v + w) - S(v) - S(w)\right) \ .$$

Hence, giving a nondegenerate quadratic form on a linear space is equivalent to giving a pseudo-inner product on that space. And: giving a pseudo-Riemannian metric β on a manifold Q is equivalent to giving a function $S : TQ \to \mathbb{R}$ of positions and velocities, for which for every $q \in Q$, the restriction $S|_{T_qQ}$ to T_qQ is a nondegenerate quadratic form. The requirement that the metric β is C^k therefore is equivalent to the requirement that $S_\beta : TQ \to \mathbb{R}$ is a C^k function.

This shows that giving a pseudo-Riemannian metric on a manifold Q is equivalent to giving a kinetic energy function S on TQ. Hence, by "pseudo-Riemannian metric" we will sometimes mean β and sometimes also the corresponding kinetic energy S.

3.2 Geodesics

Let β be a Riemannian metric (not a pseudo-Riemannian metric) on the manifold Q. If $\dot{q} \in T_qQ$, then we say that its *length* $||\dot{q}||_q$ is given by

$$||\dot{q}||_q := \sqrt{\beta_q(\dot{q}, \dot{q})} \ .$$

Now if $\gamma : [a, b] \to Q$ is a smooth curve in Q, then $\frac{d\gamma(t)}{dt} \in T_{\gamma(t)}Q$, which allows us to define the length of the curve γ as

$$l(\gamma) := \int_a^b \left|\left|\frac{d\gamma(t)}{dt}\right|\right|_{\gamma(t)} dt \ .$$

The following proposition shows that this definition is independent of the parametrization of γ:

Proposition 3.4 *Let $\psi : [c,d] \to [a,b]$ be an orientation preserving diffeomorphism, that is: $\psi(c) = a, \psi(d) = b$ and $\frac{d\psi}{ds} > 0$. Then $l(\gamma \circ \psi) = l(\gamma)$.*

Proof:

$$l(\gamma \circ \psi) = \int_c^d \left|\left| \frac{d}{ds}(\gamma \circ \psi)(s) \right|\right|_{\gamma(\psi(s))} ds = \int_c^d \left|\left| \frac{d\gamma(t)}{dt} \right|\right|_{\gamma(t)} \Bigg|_{t=\psi(s)} \cdot \frac{d\psi(s)}{ds} ds = l(\gamma) \ ,$$

where the second equality follows from the chain rule and the bilinearity of $\beta_{\gamma(t)}$ and the last equality follows from a substitution of variables $t = \psi(s)$. $\square$

If $\frac{d\gamma(t)}{dt} \neq 0$ for all t, the ambiguity in the parameterization can be removed by requiring that γ be parameterized by *arclength*. We say that a curve $\gamma : [a,b] \to Q$ is parameterized by arclength if

$$l(\gamma|_{[a,t]}) = t - a \ .$$

This is of course equivalent to $||\frac{d\gamma(t)}{dt}||_{\gamma(t)} = 1$.

Note that $l(\gamma)$ is the action integral along γ of the Lagrangian $L : TQ \to \mathbb{R}$ defined by $L(q,\dot{q}) = \sqrt{\beta_q(\dot{q},\dot{q})}$. We conclude that if $\gamma : [a,b] \to Q$ is a C^2 curve with $\gamma(a) = q_0, \gamma(b) = q_1$ is the shortest C^2 curve from q_0 to q_1, then γ must satisfy the Euler-Lagrange equation $[L]^\gamma = 0$.

Unfortunately though, the Lagrangian L is a bit nasty. First of all, it is not differentiable at the points $(q,\dot{q}) \in TQ$ for which $\dot{q} = 0$, because of the square root. A more serious problem is that L is degenerate: because $L(q,\lambda\dot{q}) = \lambda L(q,\dot{q})$ for all $\lambda > 0$, the derivative $\frac{\partial^2 L(q,\dot{q})}{\partial\dot{q}\partial\dot{q}}$ has $\dot{q}$ in its kernel. This implies that the Euler-Lagrange equations $[L]^\gamma = 0$ for L do not give rise to an explicit system of second order differential equations. Of course, this is closely related to the fact that any reparameterization of a shortest curve is also a shortest curve.

The solutions to $[L]^\gamma = 0$ become unique if we require that they are parameterized by arclength. Indeed, if γ is parameterized by arclength, then $L = 1$ along γ. Now recall the definition of the kinetic energy $T = \frac{1}{2}L^2$, i.e. $T(q,\dot{q}) = \frac{1}{2}\beta_q(\dot{q},\dot{q})$ and observe that $dT = LdL$. This implies that $[L]^\gamma = [T]^\gamma$ if γ is parameterized by arclength.

T is of course a nondegenerate Lagrangian. Moreover, as T is a constant of motion for the equations $[T]^\gamma = 0$, the solution curves of $[T]^\gamma = 0$ for which $T = \frac{1}{2}$ are automatically solutions of $[L]^\gamma = 0$ parameterized by arclength. This inspires the following definition:

Definition 3.5 *A geodesic in a pseudo-Riemannian manifold (Q,β) is a solution to the Euler-Lagrange equations*

$$[S]^q = 0 \ ,$$

where the kinetic energy $S : TQ \to \mathbb{R}$ is defined by $S : \dot{q} \in T_qQ \mapsto \frac{1}{2}\beta_q(\dot{q},\dot{q})$.

Note that in the definition of a geodesic we did not require the metric to be a Riemannian one. In a Riemannian manifold, the geodesics for which $S = \frac{1}{2}$ are those parameterized by arclength. Because S has the interpretation of kinetic energy, we also say that the geodesics describe the motion of a free particle in Q, where "free" refers to the absence of external forces.

Remark 3.6 (Unit tangent bundle) The flow of the Euler-Lagrange equations $[T]^q = 0$ on TQ is called the *geodesic flow.* Because S is a constant of motion for the geodesic flow, it leaves the so-called *unit tangent bundle*

$$(TQ)_1 := \{(q, \dot{q}) \in TQ \; | S(q, \dot{q}) = \frac{1}{2} \; \}$$

invariant. If β is a Riemannian metric, the intersection of $(TQ)_1$ with the tangent space T_qQ is diffeomorphic to a $n-1$-dimensional sphere. If moreover Q is compact, this implies that $(TQ)_1$ is a compact manifold, and hence the geodesic flow restricted to the unit tangent bundle is complete (i.e. solutions exist for all time).

One can prove that if the so-called *sectional curvatures* of a compact Riemannian manifold are negative, then the geodesic flow on $(TQ)_1$ is *mixing.* Loosely speaking, this means that the geodesic flow mixes or "stirs" the elements of $(TQ)_1$ very well.

Remark 3.7 (Einstein's general relativity) Pseudo-Riemannian metrics that are in some sense "minimally" nonpositive (never mind the exact definition now) are called Lorentzian metrics and Q is then called a Lorentzian manifold. This is the setting of Einstein's theory of general relativity, in which Q is a 4-dimensional manifold called "space-time" and on which the Lorentzian metric β describes a gravitational field.

We say that the curve $t \mapsto q(t), I \to Q$ in the Lorentzian manifold Q is *space-like* if $\beta(q(t))(dq(t)/dt, dq(t)/dt) > 0$ for all t, *time-like* if $\beta(q(t))(dq(t)/dt, dq(t)/dt) < 0$ for all t and *light-like* if $\beta(q(t))(dq(t)/dt, dq(t)/dt) = 0$ for all t.

Einstein's theory now says that a free massive relativistic particle is described by a time-like geodesic in Q for the Lorentzian metric, while light follows the light-like geodesics in space-time. No particle can travel on a space-like curves, as such a particle would move faster than light.

3.3 The geodesic equations

In local coordinates, the equations of motion for geodesics are found by writing out the Euler-Lagrange equations $[S]^q_i = 0$ for the kinetic energy

$$S(q, \dot{q}) = \frac{1}{2}\beta(q)(\dot{q}, \dot{q}) = \frac{1}{2}\sum_{i,j=1}^{n} \beta_{ij}(q)\dot{q}_i\dot{q}_j \; .$$

From the general formula (2.8) we see that these equations read $\frac{dq_l}{dt} = \dot{q}_l$ and

$$\sum_{m=1}^{n} \beta_{lm}(q)\frac{d\dot{q}_m}{dt} + \frac{1}{2}\sum_{j,k=1}^{n} \left(\partial_j\beta_{kl}(q) + \partial_k\beta_{jl}(q) - \partial_l\beta_{jk}(q)\right)\dot{q}_j\dot{q}_k \; . \tag{3.1}$$

If we now denote by

$$\beta^{lm}(q) = \left(\beta(q)^{-1}\right)_{lm}$$

the (l,m)-th element of the inverse of the matrix $\beta(q)$, then multiplying (3.1) by $\beta^{il}(q)$ and summing over l, we obtain the explicit formulas

$$\frac{dq_i}{dt} = \dot{q}_i \ , \ \frac{d\dot{q}_i}{dt} = -\sum_{j,k=1}^{n} \Gamma^i_{jk}(q)\dot{q}_j\dot{q}_k \ , \tag{3.2}$$

where

$$\Gamma^i_{jk}(q) := \frac{1}{2}\sum_{l=1}^{n} \beta^{il}(q)\left(\partial_j\beta_{kl}(q) + \partial_k\beta_{jl}(q) - \partial_l\beta_{jk}(q)\right) \ . \tag{3.3}$$

are called the *Christoffel symbols* of the metric.

The differential equations (3.2) have the following special property: if the curve $t \mapsto \gamma(t)$ solves (3.2) and $a \in \mathbb{R}$ is a constant, then the curve $t \mapsto \delta(t)$ defined by $\delta(t) := \gamma(at)$ also solves (3.2). It is straightforward to check this. A differential equation with this property is called a *spray.*

3.4 Excursion: the Jacobi metric

We will show in this section that solution curves of *natural* mechanical systems can be viewed as the geodesics of a special metric.

Let $L = S - V : TQ \to \mathbb{R}$ be a natural Lagrangian, which means that $S|_{T_qQ}$ is a nondegenerate quadratic form and V is constant on each T_qQ. In local coordinates $(q,\dot{q})$ for TQ this just means that $S(q,\dot{q}) = \frac{1}{2}\sum_{j,k=1}^n \beta_{jk}(q)\dot{q}_j\dot{q}_k$ and $V = V(q)$. In the same local coordinates, the Euler-Lagrange equations for such L read

$$\frac{d^2q_i}{dt^2} = -\sum_{j,k=1}^{n} \Gamma^i_{jk}(q)\frac{dq_j}{dt}\frac{dq_k}{dt} - \sum_{l=1}^{n} \beta^{il}(q)\frac{\partial V(q)}{\partial q_l} \ , \tag{3.4}$$

with Γ^i_{jk} as in (3.3). We saw before that the total energy $E = T + V : TQ \to \mathbb{R}$ is a constant of motion for equations (3.4). We now have the following theorem that characterizes the solution curves of equations (3.4) in terms of the geodesics of a special metric:

Theorem 3.8 (Jacobi-Maupertuis principle) *Let $S : TQ \to \mathbb{R}$ be a smooth pseudo-Riemannian metric and let $V : Q \to \mathbb{R}$ be a smooth potential energy function. Let $t \mapsto q(t), I \to Q$ be a curve in Q such that $E\left(q(t), \frac{dq(t)}{dt}\right) = e \in \mathbb{R}$ and $V(q(t)) \neq e$ for all t. Then the map $t \mapsto s(t), I \to \mathbb{R}$ defined by*

$$s(t) = 2\int_0^t e - V(q(\tau))d\tau \ .$$

is a diffeomorphism onto its image J. We denote its inverse by $s \mapsto t(s), J \to I$.

Moreover, the curve $t \mapsto q(t)$ in Q is a solution to the Euler-Lagrange equation $[S - V]^q = 0$, if and only if the curve $s \mapsto q(t(s)), J \to Q$ is a geodesic of the "Jacobi metric"

$$\tilde{S} = (e - V)S \ .$$

Proof: Because $\frac{ds(t)}{dt} = e - V(q(t)) \neq 0$, the inverse function theorem guarantees that $t \mapsto s(t)$ is a diffeomorphism onto its image. Let us denote the reparametrized curve by $s \mapsto \underline{q}(s) := q(t(s))$, or equivalently, $\underline{q}(s(t)) = q(t)$.

We work in local coordinates on Q now. Differentiation of the identity $q(t) = \underline{q}(s(t))$ with respect to t twice leads to the identities $\frac{dq_i}{dt} = 2(e - V(\underline{q}))\frac{d\underline{q}_i}{ds}$ and $\frac{d^2 q_i}{dt^2} = 4(e - V(\underline{q}))^2 \frac{d^2\underline{q}_i}{ds^2} - 4(e - V(\underline{q}))\frac{d\underline{q}_i}{ds}\sum_{l=1}^n \left.\frac{\partial V(q)}{\partial q_l}\right|_{q=\underline{q}} \cdot \frac{d\underline{q}_l}{ds}$. We conclude that the curve $t \mapsto q(t)$ solves equations (3.4) if and only if $s \mapsto \underline{q}(s)$ satisfies

$$\frac{d^2\underline{q}_i}{ds^2} = -\sum_{j,k=1}^n \Gamma^i_{jk}\frac{d\underline{q}_j}{ds}\frac{d\underline{q}_k}{ds} + \frac{1}{e-V}\frac{d\underline{q}_i}{ds}\sum_{l=1}^n \frac{\partial V}{\partial q_l}\frac{d\underline{q}_l}{ds} - \frac{1}{4\,(e-V)^2}\sum_{l=1}^n \beta^{il}\frac{\partial V}{\partial q_l} \ .$$

Using that $e - V = \frac{1}{2}\sum_{j,k=1}^n \beta_{jk}\frac{dq_j}{dt}\frac{dq_k}{dt} = 2\,(e-V)^2\sum_{j,k=1}^n \beta_{jk}\frac{d\underline{q}_j}{ds}\frac{d\underline{q}_k}{ds}$ along solutions, we then finally find that

$$\frac{d^2\underline{q}_i}{ds^2} = -\sum_{j,k=1}^n \tilde{\Gamma}^i_{jk}\frac{d\underline{q}_j}{ds}\frac{d\underline{q}_k}{ds} \ ,$$

in which

$$\tilde{\Gamma}^i_{jk} = \Gamma^i_{jk} - \frac{1}{2(e-V)}\left(\delta_{ik}\frac{\partial V}{\partial q_j} + \delta_{ij}\frac{\partial V}{\partial q_k} - \beta_{jk}\sum_{l=1}^n \beta^{il}\frac{\partial V}{\partial q_l}\right) \ . \tag{3.5}$$

Incidentally, the $\tilde{\Gamma}^i_{jk}$ are also exactly the Cristoffel symbols of the metric

$$\tilde{S}(q,\dot{q}) = (e - V(q))S(q,\dot{q})$$

defined on the subset of $q \in Q$ for which $V(q) \neq e$. This last claim is easy to verify by a short computation. □

When $E = S + V = e$, then the condition that $V \neq e$ is equivalent to the condition that $S \neq 0$. Hence the Jacobi-Maupertuis principle holds for curves that never have zero kinetic energy. When S defines a Riemannian metric, then $S(q,\dot{q}) \neq 0$ if and only if $S(q,\dot{q}) > 0$ if and only if $\dot{q} \neq 0$ and the condition that $e \neq V$ implies that $e - V > 0$. Hence the Jacobi metric $(e - V)S$ then is automatically a Riemannian metric.

3.5 Exercises

Exercise 3.1 (Geodesics on the sphere) *Denote by S^n the n-dimensional sphere:*

$$S^n := \{x \in \mathbb{R}^{n+1} \mid \langle x, x\rangle = 1\} \ .$$

Prove that for every $x \in S^n$ and every nonzero $\dot{x} \in \mathbb{R}^n$ with $\langle x, \dot{x}\rangle = 0$, the curve

$$\gamma_{x,\dot{x}}(t) := \cos(||\dot{x}||t)x + \frac{\sin(||\dot{x}||t)}{||\dot{x}||}\dot{x}$$

is a geodesic with $\gamma_{x,\dot{x}}(0) = x, \gamma'_{x,\dot{x}}(0) = \dot{x}$.

Exercise 3.2 *[The hyperbolic half plane] I copied this exercise from the lecture notes on Classical Mechanics of J.J. Duistermaat.*

Let $H = \{z \in \mathbb{C} \mid \mathrm{Im}\ z > 0\}$ be the complex upper half plane, which can be viewed as an open subset of $\mathbb{R}^2$ by identifying $x + iy \in \mathbb{C}$ with $(x, y) \in \mathbb{R}^2$. Moreover, let $SL(2, \mathbb{R})$ be the group of 2×2 matrices A with real coefficients and $\det A = 1$. For every $A \in SL(2, \mathbb{R})$, we define the fractional linear transformation $\Phi_A : \mathbb{C} \to \mathbb{C}$ by

$$\Phi_A(z) := \frac{az + b}{cz + d} \text{ if } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ .$$

Prove the following results:

- *Φ_A is complex differentiable with derivative $\Phi'_A(z) = (cz + d)^{-2}$ and maps $\Phi_A(H) \subset H$. Moreover, $\Phi_I = \mathrm{id}_{\mathbb{C}}$ and $\Phi_A \circ \Phi_B = \Phi_{AB}$. Φ_A is a diffeomorphism of H with $(\Phi_A)^{-1} = \Phi_{A^{-1}}$.*
- *For every $z \in H$ and nonzero $v \in \mathbb{C}$ there is exactly one $A \in SL(2, \mathbb{R})/\{\pm 1\}$ with $\Phi_A(z) = i$ and $\Phi'(z)v$ a positive multiple of i.*
- *Let β be a Riemannian structure on H with the property that every fractional linear transformation Φ_A is an isometry of β. Then β is uniquely defined by β_i by*

  $$\beta_z(v, v) = \beta_i(\Phi'_A(z)v, \Phi'_A(z)v) \ ,$$

  for every $A \in SL(2, \mathbb{R})$ with $\Phi_A(z) = i$.
- *This defines a Riemannian structure on H for which every fractional linear transformation is an isometry if and only if*

  $$\beta_i(v, v) = \beta_i(\Phi'_A(i)v, \Phi'_A(i)v)$$

  for all $A \in SL(2, \mathbb{R})$ with $\Phi_A(i) = i$. Equivalently, there is a $c > 0$ such that $\beta_i(v, v) = c|v|^2$. Choose $c = 1$. In this way, we get

  $$\beta_z(v, v) = (\mathrm{Im}\ z)^{-2}|v|^2 \ , \ z \in H, v \in \mathbb{C} \ .$$

- *The reflection* $S : x + iy \mapsto -x + iy$ *is an isometry of this metric. If* γ *is a geodesic with* $\gamma(0) = i$ *and* $\gamma'(0) = i$*, then* $\delta := S \circ \gamma$ *is also a geodesic with* $\delta(0) = i$ *and* $\delta'(0) = i$*. We have that* $\delta = \gamma$*, i.e.* $\gamma(t) = iy(t)$ *for some positive real valued function* $y(t)$*. The condition that the geodesic is parameterized by arclength leads to the conclusion that* $y'/y = 1$*, so that* $\gamma(t) = ie^t$*.*

- *Every geodesic parametrized by arclength is of the form* $\Phi_A \circ \gamma$ *with* γ *as above. If* $c \neq 0$ *and* $d \neq 0$ *respectively, then*

$$\lim_{t\to\infty} \delta(t) = \frac{a}{c} \text{ and } \lim_{t\to-\infty} \delta(t) = \frac{b}{d} \ .$$

If $c \neq 0$ *and* $d \neq 0$*, then the orbit of* δ *is a half circle with its center on the real axis. If* $c = 0$ *or* $d = 0$*, then the orbit of* δ *is a vertical half line.*

Remark 3.9 H is called the *hyperbolic half plane*. The *hyperbolic metric* β makes it a surface of negative "curvature". The hyperbolic half plane is a standard example of a non-Euclidean geometry.

Exercise 3.3 (Einstein's special relativity) *In Einstein's theory of special relativity, one introduces the so-called Lorentzian space-time. Let us consider the case of one space-dimension, i.e. space-time is* $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$*. Let us denote its elements by* (t, x)*, that have the interpretation of the time- and space-coordinates respectively.*

Let $c > 0$ *be a real constant, with the interpretaton of the speed of light, and let* $\alpha > 0$ *be another constant, to be determined later. For* $(t, x) \in \mathbb{R}^2$*, let us now define on the tangent space* $T_{(t,x)}\mathbb{R}^2 \cong \mathbb{R}^2$*, the bilinear form*

$$\beta_{(t,x)} : ((\dot{t}^1, \dot{x}^1), (\dot{t}^2, \dot{x}^2)) \mapsto \alpha(-c^2\dot{t}^1\dot{t}^2 + \dot{x}^1\dot{x}^2) \ .$$

- *Show that* β *defines a pseudo-Riemannian metric of index* 1*. It is called the* Lorentz metric*.*

- *Write down the geodesic equations that are defined by this metric. Show that the solution curves are straight lines. What distinguishes the time-like, light-like and space-like geodesics?*

- *Denote by* $S : T(\mathbb{R}^2) \to \mathbb{R}$ *the kinetic energy:* $S(t, x, \dot{t}, \dot{x}) = \frac{1}{2}\alpha(-c^2\dot{t}^2 + \dot{x}^2)$*. Let us parameterize a time-like geodesic by time, i.e. consider the time-like geodesic* $s \mapsto (s, sv)$*. Our experience tells us that in the "classical limit" where* $|v|$ *is small, the kinetic energy* S *should of course increase like* $\frac{1}{2}mv^2$*. Show that this implies that* $\alpha = m$ *and that* $S = -\frac{1}{2}mc^2 \left(1 - (v/c)^2\right)$*.*

- *For simplicity, assume now that the speed of light is* $c = 1$*. A linear map* $L : \mathbb{R}^2 \to \mathbb{R}^2$ *is called a* Lorentz transformation *if it is an isometry of the Lorentz metric and the collection of all Lorentz transformation is called the* Lorentz group*. Show that a* *linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ is a Lorentz transformation if and only if its matrix has the form*

$$\text{mat } L = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

with $A^2 - C^2 = 1$, $D^2 - B^2 = 1$ and $AB - CD = 0$. Show that this is true if and only if there is a $\phi \in \mathbb{R}$ such that $A = \pm\cosh\phi, B = \pm\sinh\phi, C = \pm\sinh\phi$ and $D = \pm\cosh\phi$, while at the same time sign$(AB) =$ sign(CD).

- *Assume again that $c = 1$. Show that a Lorentz transformation sends the light-cone $|t| = |x|$ to itself. Similarly, show that a Lorentz transformation sends the hyperbolas $t^2 - x^2 = E \neq 0$ to themselves.*

Exercise 3.4 *Let $S(q, \dot{q}) = \frac{1}{2}\sum_{j,k=1}^{n} \beta_{jk}(q)\dot{q}_j\dot{q}_k$. Prove that the Cristoffel symbols of the metric $\tilde{S}(q, \dot{q}) = (e - V(q))S(q, \dot{q})$ are given by (3.5)*

4 Mechanics on Lie groups

In some mechanical systems the configuration is naturally determined by an element of a Lie group G. Such mechanical systems are described by a differential equation on the tangent bundle TG of the group. The most famous example is the motion of a free rigid body, which can be viewed as the geodesic motion on $SO(3,\mathbb{R})$ with respect to a left-invariant metric.

In this chapter, we will introduce the general setting of mechanics on Lie groups and we will encounter the technique of "Euler-Poincaré reduction" that comes with it. An application is the rigid body motion. Quite remarkably, various partial differential equations for fluid flows arise in exactly the same way.

4.1 Lie groups

Let us forget about mechanics for a while. We begin by introducing groups that are at the same time differentiable manifolds. Such groups are called Lie groups:

Definition 4.1 *A C^k* Lie group *is a group G that is at the same time a C^k differentiable manifold such that*

1. *The inversion $G \to G$, $g \mapsto g^{-1}$ is a C^k map.*
2. *The multiplication $G \times G \to G$, $(g,h) \mapsto gh$ is a C^k map.*

If the group G is finite or countable, then we usually think of it as a C^∞ Lie group of dimension zero.

The most important examples of Lie groups are the matrix Lie groups, that is the subgroups of the *general linear group*

$$GL(n,\mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \mid \det A \neq 0\} \ .$$

Being an open subset of $\mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$, $GL(n,\mathbb{R})$ is an n^2-dimensional C^∞ manifold. Since matrix multiplication is a polynomial map and, by Cramer's rule, matrix inversion a rational map, $GL(n,\mathbb{R})$ is indeed a C^∞ Lie group. As we all know, the multiplication in $GL(n.\mathbb{R})$ is noncommutative.

A subgroup $H \subset G$ of a Lie group G that is at the same time a submanifold of G, is of course a Lie group itself. Examples of Lie subgroups of $GL(n,\mathbb{R})$ are the *special linear group*

$$SL(n,\mathbb{R}) = \{A \in GL(n,\mathbb{R}) \mid \det A = 1\} \ ,$$

the orthogonal group

$$O(n,\mathbb{R}) = \{A \in GL(n,\mathbb{R}) \mid AA^* = \mathrm{id}\}$$

and the *special orthogonal group*

$$SO(n,\mathbb{R}) = \{A \in GL(n,\mathbb{R}) \mid AA^* = \mathrm{id} \ , \ \det A = 1\} \ .$$

Here A^* denotes the matrix transpose of A.

4.2 A little bit of Lie theory

In this section I will present some of the theory of abstract Lie groups.

We start by defining, for every $h \in G$, the following two mappings from G to G:

$$L_h : G \to G, g \mapsto hg \ ,$$

$$R_h : G \to G, g \mapsto gh \ .$$

L_h is called 'left-translation by h' or 'left multiplication by h' and R_h is called 'right-translation by h' or 'right-multiplication by h'. Being the restriction of the multiplication map $G \times G \to G$ to $\{h\} \times G$ and $G \times \{h\}$ respectively, these maps are smooth. Furthermore, $(L_h)^{-1} = L_{h^{-1}}$ and $(R_h)^{-1} = R_{h^{-1}}$, which proves that L_h and R_h are diffeomorphisms of G.

Being a differentiable manifold, G has a tangent bundle TG consisting of the tangent spaces T_gG ($g \in G$). The tangent space at the identity element e is called the *Lie-algebra* of G, denoted

$$\mathfrak{g} := T_eG \ .$$

We remark that left-multiplication by h sends e to h: $L_h : e \mapsto h$. Hence, $T_eL_h : \mathfrak{g} \to T_hG$. Differentiation of $L_{h^{-1}} \circ L_h = L_h \circ L_{h^{-1}} = \mathrm{id_G}$ at e and h respectively shows that

$$(T_eL_h)^{-1} = T_hL_{h^{-1}} \ ,$$

i.e. $T_eL_h : \mathfrak{g} \to T_hG$ is an isomorphism.

The next step is to consider vector fields on the Lie group G. In fact, given a vector $X \in \mathfrak{g}$, we can define a vector field $v^l_X : G \to TG$ on G by

$$v^l_X(g) := T_eL_g \cdot X \in T_gG \ .$$

Proposition 4.2 *The vector field v^l_X is* left-invariant*, that is for any $h \in G$, we have*

$$(L_h)_* v^l_X = v^l_X \ .$$

Proof: $((L_h)_* v^l_X)(g) = T_{(L_h)^{-1}(g)}L_h \cdot v^l_X((L_h)^{-1}(g)) = T_{h^{-1}g}L_h \cdot v^l_X(h^{-1}g) = T_{h^{-1}g}L_h \cdot T_eL_{h^{-1}g} \cdot X = T_eL_g \cdot X = v^l_X(g)$. □

Also, a left-invariant vector field is uniquely determined by its value $v(e) \in \mathfrak{g}$ at the identity, because left-invariance implies that $v(g) = ((L_g)_* v)(g) = T_eL_g \cdot v(e)$. Thus, we have the following correspondence:

Proposition 4.3 *Let G be a C^{k+1} Lie group. Denote by $\mathcal{X}^k_l(G) \subset \mathcal{X}^k(G)$ the vector space of left-invariant C^k vector fields on G. Then the linear map*

$$X \mapsto v^l_X \ , \ \mathfrak{g} \to \mathcal{X}^k_l(G)$$

is a bijection.

Some people actually define the Lie algebra of G as the space of left-invariant vector fields on G. Finally, the above correspondence allows us to define a bracket in $\mathfrak{g}$:

Definition 4.4 *Let G be a C^k Lie group and let $X, Y \in \mathfrak{g}$. Then we define the Lie bracket $[X, Y] \in \mathfrak{g}$ as*

$$[X, Y] := -[v_X^l, v_Y^l](e) \text{ for all } X, Y \in \mathfrak{g} \ .$$

Here the right hand side is minus the Lie bracket of the vector fields v_X^l and v_Y^l on G, evaluated at the identity. The minus sign is just due to convention. This definition is designed in such a way that the following proposition holds:

Proposition 4.5 *The map $X \mapsto v_X^l$ from $\mathfrak{g}$ to $\mathcal{X}^k(G)$ is a Lie-algebra anti-homomorphism, i.e.*

$$[v_X^l, v_Y^l] = -v_{[X,Y]}^l \ .$$

Here, the bracket on the left hand side denotes the Lie bracket of $\mathcal{X}^k(G)$ and the bracket on the right hand side is the Lie bracket of $\mathfrak{g}$ defined above.

Proof: By definition, both the vector field $v_{[X,Y]}^l$ and the Lie bracket $-[v_X^l, v_Y^l]$ take the value $[X, Y]$ in e. Because the Lie bracket of two left-invariant vector fields is again left invariant, both vector fields are left invariant. Because the value at e determines a left invariant vector field uniquely, these vector fields are equal. $\square$

The proposition immediately implies that the Lie bracket on $\mathfrak{g}$ inherits the properties of the Lie bracket for vector fields: anti-symmetry and the Jacobi-identity.

In exercise 4.1 we will investigate the particular case that G is the matrix Lie group $GL(n, \mathbb{R})$. Its Lie algebra is the space of $n \times n$-matrices $\mathfrak{gl}(n, \mathbb{R})$ and it turns out that the Lie bracket is simply the matrix commutator

$$[A, B] := A \cdot B - B \cdot A \ .$$

4.3 Euler-Poincaré reduction

Suppose now that we study a mechanical system for which the configuration space is actually a Lie group. An example is the famous rigid body for which this Lie group is the rotation group $SO(3, \mathbb{R})$. We will encounter the rigid body in the next section.

In general, let G be a Lie group and assume that on the Lie algebra $\mathfrak{g}$ some smooth function $L_e : \mathfrak{g} \to \mathbb{R}$ is defined. This L_e extends to a Lagrangian L on TG by setting, for $\dot{g} \in T_g G$,

$$L(\dot{g}) := L_e(T_g L_{g^{-1}} \cdot \dot{g}) \ .$$

By construction this L is left invariant, that is $L \circ TL_h = L$ for all $h \in G$, and it is the unique left invariant Lagrangian that equals L_e on $\mathfrak{g}$. In many examples, L_e is a quadratic form on $\mathfrak{g}$ and thus, L defines a metric on G. But in the discussion that follows, L_e can be completely arbitrary.

Recall that a curve $t \mapsto \gamma(t), [a,b] \to G$ satisfies the Euler-Lagrange equations $[L]^\gamma(t) = 0$ if and only if it is stationary for the action integral

$$A(\gamma) = \int_a^b L\left(\gamma'(t)\right) dt \ ,$$

with respect to C^2 variations with fixed endpoints. But we can say more:

Theorem 4.6 (Euler-Poincaré reduction) *The following are equivalent:*

- $\gamma : [a,b] \to G$ *satisfies the Euler-Lagrange equations for the left-invariant Lagrangian* $L : TG \to \mathbb{R}$.
- *The curve* $\lambda(t) = T_{\gamma(t)} L_{\gamma(t)^{-1}} \cdot \gamma'(t), [a,b] \to \mathfrak{g}$ *is stationary for the 'reduced' action-integral*
$$a(\lambda) = \int_a^b L_e(\lambda(t))dt$$
with respect to all variations of $t \mapsto \lambda(t)$ *of the form*
$$(t,\varepsilon) \mapsto \lambda(t) + \varepsilon \left(\frac{d\delta(t)}{dt} + [\lambda(t), \delta(t)] \right) \ ,$$
with zero endpoints $\delta(a) = \delta(b) = 0$.

Proof: We start by remarking that the left-invariance of the Lagrangian implies that

$$\int_a^b L(\gamma'(t))dt = \int_a^b L_e(\lambda(t))dt \ .$$

Assume the left hand side is stationary for variations with fixed endpoints. Thus, let $(t,\varepsilon) \mapsto \tilde{\gamma}(t,\varepsilon), [a,b] \times (-\varepsilon_0, \varepsilon_0) \to G$ be a variation of $t \mapsto \gamma(t) = \tilde{\gamma}(t,0)$ with fixed endpoints, i.e. $\tilde{\gamma}(a,\varepsilon) = \tilde{\gamma}(a,0)$ and $\tilde{\gamma}(b,\varepsilon) = \tilde{\gamma}(b,0)$ for all ε. Then $\tilde{\lambda} : [a,b] \times (-\varepsilon_0, \varepsilon_0) \to \mathfrak{g}$ defined by $\tilde{\lambda}(t,\varepsilon) := T_{\tilde{\gamma}(t,\varepsilon)} L_{\tilde{\gamma}(t,\varepsilon)^{-1}} \cdot \gamma'(t,\varepsilon)$ defines a variation of $t \mapsto \lambda(t) = \tilde{\lambda}(t,0)$. Define

$$\delta(t) := T_{\gamma(t)} L_{\gamma(t)^{-1}} \cdot \frac{\partial \tilde{\gamma}(t,0)}{\partial \varepsilon} \in \mathfrak{g} \ .$$

Then $\delta(a) = \delta(b) = 0$ because $\tilde{\gamma}$ has fixed endpoints. I claim that

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \tilde{\lambda}(t,\varepsilon) = \delta'(t) + [\lambda(t), \delta(t)] \ . \tag{4.1}$$

To prove formula (4.1), let us pretend that G is a matrix Lie group. It turns out that formula (4.1) is also true in the case that G is a general abstract Lie group, where the bracket is the Lie bracket of $\mathfrak{g}$. The proof of this general result is a little cumbersome though, and I will skip it in these notes.

Anyway, when G is a matrix Lie group, then $\tilde{\lambda}(t,\varepsilon) = \tilde{\gamma}(t,\varepsilon)^{-1} \cdot \frac{\partial\gamma(t,\varepsilon)}{\partial t}$ and $\delta(t) = \tilde{\gamma}(t,0)^{-1} \cdot \frac{\partial\gamma(t,0)}{\partial\varepsilon}$. We hence find that

$$\frac{\partial\tilde{\lambda}(t,\varepsilon)}{\partial\varepsilon} = \frac{\partial}{\partial\varepsilon}\left(\tilde{\gamma}(t,\varepsilon)^{-1}\right) \cdot \frac{\partial\tilde{\gamma}(t,\varepsilon)}{\partial t} + \tilde{\gamma}(t,\varepsilon)^{-1}\frac{\partial^2\tilde{\gamma}(t,\varepsilon)}{\partial\varepsilon\partial t}$$

and

$$\delta'(t) = \frac{\partial}{\partial t}\left(\tilde{\gamma}(t,0)^{-1}\right) \cdot \frac{\partial\tilde{\gamma}(t,0)}{\partial\varepsilon} + \tilde{\gamma}(t,0)^{-1}\frac{\partial^2\tilde{\gamma}(t,0)}{\partial t\partial\varepsilon} \ .$$

Formula (4.1) now follows from consecutively evaluating these two identities in $\varepsilon = 0$, subtracting them and using the theorem for interchanging the order of differentiation. Moreover, one also needs to know that $\frac{\partial}{\partial t}\left(\gamma(t,\varepsilon)^{-1}\right) = -\gamma(t,\varepsilon)^{-1} \cdot \frac{\partial\gamma(t,\varepsilon)}{\partial t} \cdot \gamma(t,\varepsilon)^{-1}$ (and similarly for the derivative with respect to ε) and that $[\lambda,\delta] = \lambda\cdot\delta - \delta\cdot\lambda$.

The left-invariance of L means that

$$A(\tilde{\gamma}(\cdot,\varepsilon)) = \int_a^b L(\tilde{\gamma}'(t,\varepsilon))dt = \int_a^b L_e(\tilde{\lambda}(t,\varepsilon))dt = a(\tilde{\lambda}(\cdot,\varepsilon)) \ .$$

Thus, if a is stationary at λ with respect to variations of the form $\delta' + [\lambda,\delta]$, then A is stationary with respect to variations with fixed endpoints.

Also, one can produce the variation $\lambda + \varepsilon\,(\delta' + [\lambda,\delta])$ of λ by choosing a variation $\tilde{\gamma}$ to γ for which $\frac{\partial\tilde{\gamma}(t,0)}{\partial\varepsilon} = T_{\gamma(t)}L_{\gamma(t)^{-1}} \cdot \delta(t)$, which implies that if γ is stationary for A, then so is λ for a with respect to the allowed variations. $\square$

The process of reducing the Euler-Lagrange equations on $TG \cong G \times \mathfrak{g}$ to the 'Euler-Poincaré equations' on $\mathfrak{g}$ is called Euler-Poincaré reduction, because $\mathfrak{g}$ has only half the dimension of TG. Once one has found a solution curve $t \mapsto \lambda(t) \in \mathfrak{g}$, one may try to *reconstruct* the solution curve γ in G. This is done by writing

$$\frac{dg(t)}{dt} = T_eL_{g(t)} \cdot \lambda(t) \ ,$$

which is a nonautonomous, first order differential equation on G. Given an initial condition $g(0) = g_0$, its solutions are unique.

4.4 The rigid body

The configuration space of a free rigid body is the 3-dimensional Lie group

$$SO(3,\mathbb{R}) = \{A : \mathbb{R}^3 \to \mathbb{R}^3 \mid A^*A = I \ , \det A = 1\} \ ,$$

of rotations of 3-dimensional space, as the configuration of the body is determined by a rotation matrix that describes the orientation of the body with respect to a reference configuration.

How to determine the equations of motion for the rigid body in the absence of external forces such as gravity? Because the motion of the rigid body is supposed to be free, the Lagrangian of the body is the kinetic energy. Suppose now that $t \mapsto A(t) \in SO(3,\mathbb{R})$ is the motion of the body and that $A(t_0) = A_0$ is the configuration at $t = t_0$ with respect to the reference configuration. We denote by $\dot{A}(t_0) = \left.\frac{dA(t)}{dt}\right|_{t=t_0} \in T_A SO(3,\mathbb{R})$ the velocity of the body at $t = t_0$. Then we can argue that the curve $B(t) = A_0^{-1}A(t)$, with $B(t_0) = I$ and $\dot{B}(t_0) = A_0^{-1}\dot{A}(t_0) \in \mathfrak{so}(3,\mathbb{R})$, must have the same kinetic energy as the curve $A(t)$.

This shows that the Lagrangian of the rigid body is determined completely by choosing a quadratic form S_I on the Lie algebra

$$\mathfrak{so}(3,\mathbb{R}) = T_I SO(3,\mathbb{R}) = \{X \in \mathbb{R}^{3\times 3} \mid X^* + X = 0\}$$

and extending it to a kinetic energy S on $TSO(3,\mathbb{R})$ (or equivalently: a Riemannian metric on $SO(3,\mathbb{R})$) by the formula

$$S(A,\dot{A}) = S_I(A^{-1}\dot{A}) \ .$$

We observe that, by construction, this kinetic energy is a *left-invariant* Lagrangian function on $TSO(3,\mathbb{R})$.

According to the previous section, the curve $t \mapsto A(t)$ is a geodesic in $SO(3,\mathbb{R})$ for the left-invariant kinetic energy S if and only if the curve $t \mapsto \Omega(t) := A^{-1}(t) \cdot \frac{dA(t)}{dt}$ in $\mathfrak{so}(3,\mathbb{R})$ is stationary for the action integral

$$\int_a^b S_e(\Omega(t))dt$$

with respect to all variations of $t \mapsto \Omega(t)$ of the form $t \mapsto \Omega(t) + \varepsilon\,(\Xi'(t) + [\Omega(t),\Xi(t)])$ with $\Xi(a) = \Xi(b) = 0$, i.e.

$$0 = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_a^b S_e(\Omega(t) + \varepsilon(\Xi'(t) + [\Omega(t),\Xi(t)]))dt = \int_a^b \beta_e(\Omega(t),\Xi'(t) + [\Omega(t),\Xi(t)])dt \tag{4.2}$$

for all curves $t \mapsto \Xi(t)$ in $\mathfrak{so}(3,\mathbb{R})$ with fixed endpoints.

Let us be more concrete and choose a particular quadratic form on $\mathfrak{so}(3,\mathbb{R})$. For this purpose, we first of all note that an arbitrary $\Omega \in \mathfrak{so}(3,\mathbb{R})$ can be written as

$$\Omega = \begin{pmatrix} 0 & \omega_1 & \omega_2 \\ -\omega_1 & 0 & \omega_3 \\ -\omega_2 & -\omega_3 & 0 \end{pmatrix} \ .$$

So we can represent the anti-symmetric 3×3-matrix Ω as a 3-vector ω. In terms of this representation, a quadratic form on $\mathfrak{so}(3,\mathbb{R})$ is given by the formula

$$S_e(\Omega) = \frac{1}{2}\langle \omega, I\omega \rangle \ .$$

The symmetric 3×3-matrix above is called the "moments of inertia tensor" of the rigid body.

Then with a bit of work (one partial integration and a couple of identities for the inner and cross product) one finds that the right hand side of 4.2 is equal to

$$\int_a^b \langle -I\frac{d\omega(t)}{dt} + \omega(t) \times I\omega(t), \xi(t)\rangle dt \ .$$

In other words, the Euler-Poincaré equations for the free rigid body are given by the ordinary differential equation

$$\frac{d}{dt} I\omega = \omega \times I\omega \ .$$

4.5 Eulerian fluid equations

Some important fluid dynamical equations can be written in Euler-Poincaré form. The (Lie) group under consideration here is the group $Diff^\infty(X)$ of C^∞ diffeomorphisms of a bounded open subset $X \subset \mathbb{R}^n$ with C^∞ boundary ∂X. X has the interpretation of a fluid container. An element $x \in X$ has the interpretation of a fluid particle's reference position or 'fluid label'. Each diffeomorphism $\phi \in Diff^\infty(X)$ then represents a possible fluid configuration, where $\phi(x)$ has the interpretation of the position of the fluid element with fluid label x. The requirement that ϕ be a diffeomorphism is to prevent the fluid from developing shocks, particle collapse and other kinds of singularities.

We will say that $u : [a, b] \to Diff^\infty(X), t \mapsto u(t, \cdot)$ is a C^∞ curve of diffeomorphisms of X if u is C^∞ as a map from $[a, b] \times X$ to X. Such a curve describes a possible fluid motion, so that $t \mapsto u(t, x), [a, b] \to X$ describes the trajectory of the fluid element with label x. One often also requires that the fluid motions leave the volume-form $d_n x$ on X invariant, that is that the fluid is incompressible. In this case, the fluid motions are curves in $SDiff^\infty(X)$, the (Lie) group of volume preserving diffeomorphisms of X.

As usual, the Lagrangian fluid dynamical equations are obtained from Hamilton's principle: given a Lagrangian function $L = L(u, \dot{u})$, defined for diffeomorphisms $u : X \to X$ and velocities $\dot{u} := \frac{\partial u(t,\cdot)}{\partial t} : X \to \mathbb{R}^n$, it is postulated that $t \mapsto u(t)$ is a physical fluid flow if and only if

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_a^b L\left(\tilde{u}(t, \cdot, \varepsilon), \frac{\partial \tilde{u}(t, \cdot, \varepsilon)}{\partial t}\right) dt = 0$$

for every C^∞ variation $\tilde{u} : [a, b] \times (-\varepsilon_0, \varepsilon_0) \to (S)Diff^\infty(X)$ of u (i.e. $\tilde{u}(t, x, 0) = u(t, x)$) with fixed endpoints $\tilde{u}(a, x, \varepsilon) = u(a, x)$, $\tilde{u}(b, x, \varepsilon) = u(b, x)$.

Usually $L(u, \dot{u})$ is the integral over X of some density depending on u, $\dot{u}$ and their x-derivatives, i.e.

$$L(u, \dot{u}) = \int_X l(u(x), Du(x), \ldots, D^k u(x); \dot{u}(x), D\dot{u}(x), \ldots, D^m \dot{u}(x)) \ d_n x \ ,$$

but in more exotic applications L can involve integrals over the boundary of X, for instance if surface tension is taken into account. The simplest possible Lagrangians do not depend on the $x-$derivatives of u and $\dot{u}$, i.e. $L(u,\dot{u}) = \int_X l(u,\dot{u}) d_n x$. Such fluids are sometimes called nonviscous or 'ideal'. An example is $L(u,\dot{u}) = \int_X \frac{1}{2}\sum_{i=1}^n (\dot{u}_i(x))^2 d_n x$, the total kinetic energy of the fluid.

Let's note that an ideal Lagrangian has a remarkable symmetry: if $\phi : X \to X$ is any volume-preserving diffeomorphism, then

$$L(u \circ \phi^{-1}, \dot{u} \circ \phi^{-1}) = \int_X l(u(\phi^{-1}(x)), \dot{u}(\phi^{-1}(x))) d_n x = \int_X l(u(\tilde{x}), \dot{u}(\tilde{x})) d_n \tilde{x} = L(u,\dot{u}) \ .$$

This simply follows from the substitution of variables $x = \phi(\tilde{x})$, using that $\det\left(\frac{\partial \phi(\tilde{x})}{\partial \tilde{x}}\right) = 1$ if ϕ is volume-preserving. This means that the ideal Lagrangian is right-invariant, i.e. invariant under the action $(u,\dot{u}) \mapsto (u \circ \phi^{-1}, \dot{u} \circ \phi^{-1})$ of $SDiff^\infty(X)$ by right multiplication. This symmetry is called the *relabeling symmetry* of an ideal fluid and it expresses that fluid elements can be given another name or label according to any volume-preserving diffeomorphism ϕ^{-1} without changing the value of the Lagrangian. If u is itself volume-preserving, then $L(u,\dot{u}) = L(\mathrm{id}, \dot{u} \circ u^{-1}) =: L_{\mathrm{id}}(\dot{u} \circ u^{-1})$.

We shall now sketch the derivation of the Euler-Poincaré equations for a right-invariant Lagrangian on $SDiff^\infty(X)$ from a variational principle, and in particular we shall derive the Euler equations for an ideal incompressible fluid.

Instead of deriving the Euler-Lagrange equations for u, we shall exploit the right-invariance of the Lagrangian L to derive Euler-Poincaré equations for the curve of Eulerian velocity fields $\lambda : [a,b] \times X \to \mathbb{R}^n$ defined as

$$\lambda = \text{“}\ \frac{\partial u}{\partial t} \circ u^{-1}\ \text{”} : (t,x) \mapsto \left.\frac{\partial u(t,\tilde{x})}{\partial t}\right|_{\tilde{x} = u(t,\cdot)^{-1}(x)} ,$$

or implicitly:

$$\lambda(t, u(t,x)) = \frac{\partial u(t,x)}{\partial t} \ .$$

$\lambda(t,x)$ simply has the interpretation of the velocity of the fluid element that is at position x at time t. Let us discuss some properties of λ. First of all, when $x \in \partial X$, then $u(t,x) \in \partial X$ for each t, whence $\lambda(t,u(t,x)) = \frac{\partial u(t,x)}{\partial t} \in T_{u(t,x)}\partial X$, so that we can conclude that λ is tangent to ∂X.

The second property follows from differentiating the identity $\lambda_j(t,u(t,x)) = \frac{\partial u_j(t,x)}{\partial t}$ with respect to x_k to obtain

$$\frac{\partial}{\partial t}\frac{\partial u_i(x,t)}{\partial x_k} = \sum_{l=1}^n \frac{\partial \lambda_i(t,u(x,t))}{\partial x_l}\frac{\partial u_l(t,x)}{\partial x_k} ,$$

i.e. $\frac{\partial}{\partial t}\frac{\partial u(x,t)}{\partial x} = \frac{\partial \lambda(t,u(x,t))}{\partial x}\frac{\partial u(t,x)}{\partial x}$. This and the fact that $\det \frac{\partial u(t,x)}{\partial x} = 1$ identically, leads to the conclusion that

$$0 = \frac{\partial}{\partial t} \det\left(\frac{\partial u(t,x)}{\partial x}\right) = \left.\frac{d}{dh}\right|_{h=0} \det\left(\frac{\partial u(t+h,x)}{\partial x}\right) =$$
$$\left.\frac{d}{dh}\right|_{h=0} \det\left(\frac{\partial u(t+h,x)}{\partial x}\left(\frac{\partial u(t,x)}{\partial x}\right)^{-1}\right) \det\left(\frac{\partial u(t,x)}{\partial x}\right) = \mathrm{tr}\left(\frac{\partial \lambda(t,u(t,x))}{\partial x}\right) ,$$

where we have used that $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \det(I+\varepsilon E) = \mathrm{tr}(E)$. We conclude that for each t, the vector field $x \mapsto \lambda(t,x)$ on X is divergence-free: $\mathrm{div}(\lambda) = \sum_{j=1}^n \frac{\partial \lambda_j}{\partial x_j} = 0$.

After these preparations, let $u : [a,b] \to SDiff^\infty(X)$ be a C^∞ curve of volume preserving diffeomorphisms and let $\tilde{u} : [a,b] \times (-\varepsilon_0, \varepsilon_0) \to SDiff^\infty(X)$ be a C^∞ variation of u with fixed endpoints, i.e. $\tilde{u}(t,x,0) = u(t,x)$, $\tilde{u}(a,x,\varepsilon) = u(a,x)$ and $\tilde{u}(b,x,\varepsilon) = u(b,x)$. Define $\tilde{\lambda}, \tilde{\delta} : [a,b] \times X \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^n$ implicitly by

$$\tilde{\lambda}(t, \tilde{u}(t,x,\varepsilon), \varepsilon) = \frac{\partial \tilde{u}(t,x,\varepsilon)}{\partial t} , \tag{4.3}$$
$$\tilde{\delta}(t, \tilde{u}(t,x,\varepsilon), \varepsilon) = \frac{\partial \tilde{u}(t,x,\varepsilon)}{\partial \varepsilon} , \tag{4.4}$$

and set $\delta(t,x) := \tilde{\delta}(t,x,0)$. Note that $\lambda(t,x) = \tilde{\lambda}(t,x,0)$ and that $\tilde{\delta}(a,x,\varepsilon) = \tilde{\delta}(b,x,\varepsilon) = 0$ because $\tilde{u}$ has fixed endpoints. Moreover, a similar argument as above shows that $x \mapsto \tilde{\lambda}(t,x,\varepsilon)$ and $x \mapsto \tilde{\delta}(t,x,\varepsilon)$ are tangent to ∂X and divergence-free.

Our main observation now is that differentiating (4.3) with respect to ε and (4.4) with respect to t, evaluating the resulting equations in $(t,x,\varepsilon) = (t, \tilde{u}(t,\cdot,0)^{-1}(\tilde{x}), 0)$ and using that $\frac{\partial^2 \tilde{u}_i}{\partial t \partial \varepsilon} = \frac{\partial^2 \tilde{u}_i}{\partial \varepsilon \partial t}$, we find (please perform this computation if you don't believe it!):

$$\left.\frac{\partial \tilde{\lambda}_i(t,x,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = \frac{\partial \delta_i(t,x)}{\partial t} + \sum_{j=1}^n \left(\frac{\partial \delta_i(t,x)}{\partial x_j}\lambda_j(t,x) - \frac{\partial \lambda_i(t,x)}{\partial x_j}\delta_j(t,x)\right) .$$

This proves the "only if" part of:

Theorem 4.7 (Euler-Poincaré for incompressible ideal fluids) *Let $L = L(u,\dot{u}) = L_{\mathrm{id}}(\dot{u} \circ u^{-1})$ be a right invariant Lagrangian function, defined for volume preserving diffeomorphisms $u : X \to X$ and vector fields $\dot{u} : X \to \mathbb{R}^n$. Then the following are equivalent:*

- *The curve $t \mapsto u(t)$ in $SDiff^\infty(X)$ is stationary for the action integral $\int_a^b L(u(t), \frac{\partial u(t)}{\partial t})dt$ with respect to volume-preserving variations with fixed endpoints.*
- *The curve of divergence-free vector fields $t \mapsto \lambda(t) := \frac{\partial u(t)}{\partial t} \circ u(t)^{-1}$ is stationary for the action integral $\int_a^b L_{\mathrm{id}}(\lambda(t))dt$ with respect to variations $\tilde{\lambda}$ of λ of the form*

$$\tilde{\lambda} = \lambda + \varepsilon\left(\frac{\partial \delta}{\partial t} + \sum_{j=1}^n \left(\frac{\partial \delta}{\partial x_j}\lambda_j - \frac{\partial \lambda}{\partial x_j}\delta_j\right)\right) ,$$

such that δ is tangent to ∂X, divergence-free and $\delta(a,x) = \delta(b,x) = 0$.

I leave the "if" part of this theorem as a (rather difficult) exercise.

As an example, let us compute the Euler equations for an ideal fluid, that is the Euler-Poincaré equations for the kinetic Lagrangian

$$L(u,\dot{u}) = ||\dot{u}||^2_{L_2} := \int_X \frac{1}{2}\sum_{i=1}^n (\dot{u}_i(x))^2 d_n x = \int_X \frac{1}{2}\sum_{i=1}^n ((\dot{u}_i \circ u^{-1})(x))^2 d_n x \ .$$

According to Theorem 4.7, for all curves $(t,x)\mapsto \delta(t,x), [a,b]\times X \to \mathbb{R}^n$ with $\delta(t,\cdot)$ tangent to ∂X, $\mathrm{div}(\delta(t,\cdot)) = 0$ and $\delta(a,x) = \delta(b,x) = 0$, the curve $t\mapsto \lambda(t) := \frac{\partial u(t)}{\partial t}\circ u(t)^{-1}$ should then satisfy

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\int_a^b\int_X \frac{1}{2}\sum_{i=1}^n \left(\lambda_i(x) + \varepsilon\left[\frac{\partial \delta_i}{\partial t} + \sum_{j=1}^n\left(\frac{\partial\delta_i}{\partial x_j}\lambda_j - \frac{\partial \lambda_i}{\partial x_j}\delta_j\right)\right]\right)^2 dx = 0 \ .$$

Integration by parts yields that the left hand side of this expression is equal to

$$\int_a^b\int_X\sum_{i=1}^n\lambda_i\left(\frac{\partial\delta_i}{\partial t} + \sum_{j=1}^n\left(\frac{\partial\delta_i}{\partial x_j}\lambda_j - \frac{\partial\lambda_i}{\partial x_j}\delta_j\right)\right)dxdt = \tag{4.5}$$

$$\int_a^b\int_X\sum_{i=1}^n -\delta_i\left(\frac{\partial\lambda_i}{\partial t} + \sum_{j=1}^n\lambda_j\frac{\partial\lambda_i}{\partial x_j} + \lambda_i\mathrm{div}(\lambda)\right) + \left(\frac{1}{2}\sum_{j=1}^n\lambda_j^2\right)\mathrm{div}(\delta)\ dxdt =$$

$$\int_a^b\int_X\sum_{i=1}^n -\delta_i\left(\frac{\partial\lambda_i}{\partial t} + \sum_{j=1}^n\lambda_j\frac{\partial\lambda_i}{\partial x_j}\right)dxdt \ .$$

If $X\subset\mathbb{R}^n$ is bounded, then the space of vector fields on X tangent to ∂X is the L_2-orthogonal sum of the divergence free vector fields and the gradient vector fields. This is a consequence of Hodge-de Rham theory and unfortunately it would go too far to explain this in detail. Nevertheless, we have derived the Euler equations for an ideal incompressible fluid:

$$\frac{\partial\lambda_i}{\partial t} + \sum_{j=1}^n\lambda_j\frac{\partial\lambda_i}{\partial x_j} = -\frac{\partial p}{\partial x_i}\ ,\ \mathrm{div}(\lambda) = 0\ .$$

The function p is implicitly defined by the condition $\mathrm{div}(\lambda) = 0$ and is called the *pressure* of the fluid.

The above derivation of Euler's equations for an incompressible fluid as the Euler-Poincaré equations for the geodesics on the special diffeomorphism group with a right-invariant metric, is due to Arnol'd.

4.6 Exercises

Exercise 4.1 (Matrix Lie groups) *The collection of all invertible $n\times n$ matrices is called the n-th general linear group:*

$$GL(n,\mathbb{R}) := \{A\in\mathbb{R}^{n\times n} \mid \det A\neq 0\}\ .$$

Prove that:

- *$GL(n,\mathbb{R})$ is an open subset of $\mathbb{R}^{n\times n}$ and hence a manifold of dimension n^2.*
- *$GL(n,\mathbb{R})$ is a group under matrix multiplication.*
- *$GL(n,\mathbb{R})$ is a C^∞ Lie group, that is the multiplication and inversion are C^∞ maps.*
- *For $A \in GL(n,\mathbb{R})$, the tangent space $T_A GL(n,\mathbb{R})$ is isomorphic to $\mathbb{R}^{n\times n}$. Remark: The Lie algebra $T_I GL(n,\mathbb{R})$ is denoted $\mathfrak{gl}(n,\mathbb{R})$.*
- *For $A \in GL(n,\mathbb{R})$, left translation over A is the map $L_A : B \mapsto AB$.*
- *The tangent map $T_I L_A$ sends $E \in \mathfrak{gl}(n,\mathbb{R})$ to $AE \in T_A GL(n,\mathbb{R})$.*
- *For $E \in \mathfrak{gl}(n,\mathbb{R})$, the unique left-invariant vector field v_E^l on $GL(n,\mathbb{R})$ which takes the value $v_E^l(I) = E$, is given by $v_E^l(A) = AE$.*
- *The curve $t \mapsto A(t)$ in $GL(n,\mathbb{R})$ is an integral curve of v_E^l, if and only if $\frac{dA(t)}{dt} = A(t)\cdot E$.*
- *The integral curves $t \mapsto A(t)$ of the left invariant vector field v_E^l are given by*
$$A(t) = A(0)\cdot e^{tE} \ .$$
- *Let $E_1, E_2 \in \mathfrak{gl}(n,\mathbb{R})$. The Lie bracket $[E_1,E_2] := -[v_{E_1}^l, v_{E_2}^l](I)$ is given by the commutator*
$$[E_1,E_2] = E_1\cdot E_2 - E_2\cdot E_1 \ .$$

Exercise 4.2 (The Lie algebra $\mathfrak{so}(3,\mathbb{R})$) *The 3-dimensional special orthogonal group is defined as*
$$SO(3,\mathbb{R}) := \{A \in \mathbb{R}^{3\times 3} \mid A\cdot A^* = I \ , \ \det A = 1\},$$
where A^* denotes the transpose of A. You may assume without proof that $SO(3,\mathbb{R})$ is a 3-dimensional Lie subgroup of the 9-dimensional Lie group $GL(3,\mathbb{R})$.

- *Show that the Lie algebra $\mathfrak{so}(3,\mathbb{R}) := T_I SO(3,\mathbb{R})$ consists of the skew-symmetric matrices:*
$$\mathfrak{so}(3,\mathbb{R}) := \{E \in \mathbb{R}^{3\times 3} \mid E + E^* = 0 \ \} \ .$$
- *Show that the mapping*
$$\sigma : \begin{pmatrix} 0 & e_1 & e_2 \\ -e_1 & 0 & e_3 \\ -e_2 & -e_3 & 0 \end{pmatrix} \mapsto \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \ , \ \mathfrak{so}(3,\mathbb{R}) \to \mathbb{R}^3$$
is a linear isomorphism.

- *Prove that*
$$\sigma([E_1, E_2]) = \sigma(E_2) \times \sigma(E_1) \ ,$$
where $[E_1, E_2] := E_1 \cdot E_2 - E_2 \cdot E_1$ *is the matrix commutator and* $a \times b \in \mathbb{R}^3$ *denotes the usual cross product of* 3*-vectors, that is*
$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \ .$$

Remark: We say that σ *is a Lie algebra anti-homomorphism from* $\mathfrak{so}(3, \mathbb{R})$ *with the matrix commutator to* $\mathbb{R}^3$ *with the cross product.*

Exercise 4.3 (The rigid body dynamics) *Recall Euler's equations for the rigid body*
$$I \frac{d\omega}{dt} = \omega \times I\omega \ , \tag{4.6}$$
in which the inertia matrix I *is symmetric and nondegenerate. Using the explicit Euler equations for the rigid body (**??**), show that the energy* $E = \frac{1}{2}\langle I\omega, \omega \rangle$ *is a constant of motion.*

Can you give another argument why E *is a constant of motion? Hint: View* E *as a function on* $TSO(3, \mathbb{R})$.

Prove that the solutions of the rigid body equations lie on ellipsoids. Prove that they stay bounded and are defined for all time.

Show moreover that the function $J(\omega) = \langle I\omega, I\omega \rangle$ *is conserved. Can you draw the joint level sets of* E *and* J *in* $\mathbb{R}^3$*?*

Exercise 4.4 (The hyperbolic half plane again) *In this exercise we will again study the hyperbolic half plane* $H = \{z \in \mathbb{C} \mid \mathrm{Im}\ z > 0\}$ *and we will give it the structure of a noncommutative Lie group.*

- *For* $x, y \in \mathbb{R}$, $y > 0$, *let* $\phi_{x,y} : \mathbb{R} \to \mathbb{R}$ *be the affine mapping defined by*
$$\phi_{x,y}(s) = x + ys \ .$$
Show that $\phi_{x,y} \circ \phi_{\tilde{x},\tilde{y}} = \phi_{x + y\tilde{x}, y\tilde{y}}$. *In other words, the collection of affine mappings*
$$G = \{\phi_{x,y} \mid x, y \in \mathbb{R}, y > 0\}$$
is closed under composition. Show that G *is a group under composition. Show that the identity element of* G *is* $\phi_{0,1}$ *and that* $\phi_{x,y}^{-1} = \phi_{-\frac{x}{y}, \frac{1}{y}}$.

- *Prove that the mapping* $x + iy \mapsto \phi_{x,y}$ *from* H *to* G *is bijective. Prove that the unique group operation* $* : H \times H \to H$ *that makes this mapping a group isomorphism, is given by*
$$(x + iy) * (\tilde{x} + i\tilde{y}) = (x + y\tilde{x}) + iy\tilde{y} \ .$$
From now on, $*$ *will be "the" multiplication on* H *and we will forget about the ordinary multiplication of complex numbers.*

- *Prove that* $*$ *makes* H *into an analytic Lie group. Show that the identity element of* H *is* i *and that* $(x+iy)^{-1} = \frac{-x}{y} + i\frac{1}{y}$.

- *Write* $z = x+iy$. *Show that* $L_{z^{-1}} : H \to H$ *sends* $\tilde{x}+i\tilde{y}$ *to* $\frac{\tilde{x}-x}{y} + i\frac{\tilde{y}}{y}$. *Now show that* $T_z L_{z^{-1}}$ *maps* $v = v_x + iv_y \in T_z H \cong \mathbb{C}$ *to* $\frac{1}{y}v := \frac{v_x}{y} + i\frac{v_y}{y} \in T_i H \cong \mathbb{C}$.

- *On* $T_i H$ *let us define an inner product by*

$$\langle v, v \rangle := |v|^2 \text{ for } v \in T_i H \cong \mathbb{C} \ .$$

Prove that the unique left-invariant metric β *on* H *with the property that* $\beta_i(v,v) = \langle v, v\rangle$ *for all* $v \in T_i H$, *is given by the hyperbolic metric of Exercise 3.2, that is*

$$\beta_z(v,v) = \frac{1}{(\mathrm{Im}\ z)^2}|v|^2 \ .$$

- *For* $\xi = \xi_x + i\xi_y \in T_i H \cong \mathbb{C}$, *show that the unique left invariant vector field* v_l^{ξ} *on* H *is given by* $v_l^{\xi}(z) = (\mathrm{Im}\ z)\xi$. *Show that the Lie bracket in the Lie algebra* $T_i H$ *is given by*

$$[\xi, \eta] := \eta_x \xi_y - \eta_y \xi_x \in \mathbb{R} \subset \mathbb{C} \ .$$

- *Let* $t \mapsto z(t) \in H$ *be a geodesic for the hyperbolic metric and write* $z(t) = x(t) + iy(t)$. *Show that* $\lambda(t) := T_{z(t)} L_{z(t)^{-1}} \cdot \frac{dz(t)}{dt} = \frac{1}{y(t)}\left(\frac{dx(t)}{dt} + i\frac{dy(t)}{dt}\right)$ *is a solution of the Euler-Poincaré equations*

$$\frac{d\lambda_x}{dt} = -\lambda_x \lambda_y \ , \ \frac{d\lambda_y}{dt} = \lambda_x^2 \ .$$

Show that the energy $\frac{1}{2}|\lambda|^2$ *is a constant of motion for these equations. Draw the phase portrait.*

Remark 4.8 One can repeat this procedure and define for $x \in \mathbb{R}^n$ and $y > 0$ the mappings $\phi_{x,y} : \mathbb{R}^n \to \mathbb{R}^n$, thus producing a "hyperbolic half-$(n+1)$-space".

Exercise 4.5 (Ideal compressible fluids) *Let* $X \subset \mathbb{R}^n$ *be a bounded open subset. For an "ideal compressible fluid" the Lagrangian function is the total kinetic energy*

$$L(u, \dot{u}) = \int_X \frac{1}{2}\sum_{i=1}^{n} (\dot{u}_i(x))^2 d_n x \ .$$

This Lagrangian is not right-invariant under the group $Diff^{\infty}(X)$. *Why? For a* C^{∞} *family of diffeomorphisms* $(t,x) \mapsto u(t,x)$, $[a,b] \times X \to X$ *show that the following are equivalent:*

- u *is stationary for the action integral* $\int_a^b L\left(u(t), \frac{\partial u(t)}{\partial t}\right) dt$ *with respect to variations with fixed endpoints.*

- *$(t,x) \mapsto u(t,x)$ solves the Euler-Lagrange equations $\frac{\partial^2 u_i(t,x)}{\partial t^2} = 0$ for all $i = 1, \ldots, n$.*

- *For each $x \in X$, the curve $t \mapsto u(t,x)$, $[a,b] \to X$ is a solution to the second order ordinary differential equation $\frac{\partial^2 u_i(t)}{\partial t^2} = 0$ ($i = 1, \ldots, n$).*

- *For each $x \in X$, $t \mapsto u(t,x)$, $[a,b] \to X$ is a geodesic in X for the Euclidean metric.*

- *$u_i(t,x) = u_i(0,x) + \left.\frac{\partial u_i(t,x)}{\partial t}\right|_{t=0} \cdot t$ for all $i = 1, \ldots, n$.*

- *$\lambda := \frac{\partial u}{\partial t} \circ u^{-1}$ (i.e. $\lambda_i(t, u(t,x)) = \frac{\partial u_i(t,x)}{\partial t}$ for all $i = 1, \ldots, n$) satisfies the Euler equation for an ideal compressible fluid*

$$\frac{\partial \lambda_i}{\partial t} + \sum_{j=1}^{n} \lambda_j \frac{\partial \lambda_i}{\partial x_j} = 0 \text{ for all } i = 1, \ldots, n \; .$$

Remark 4.9 (Burgers' equation) For $n = 1$ the Euler equation for an ideal compressible fluid is called Burgers' equation: $\frac{\partial \lambda}{\partial t} + \lambda \frac{\partial \lambda}{\partial x} = 0$.

Exercise 4.6 (1-dimensional EPDiff) *Let $X = (\alpha, \beta) \subset \mathbb{R}$. We shall study the Euler-Poincaré equation on the group $Diff^\infty(X)$ of diffeomorphisms of X as follows. Let $l : \mathbb{R}^{k+1} \to \mathbb{R}$ be a C^∞ function. For a diffeomorphism $u : X \to X$ and vector field $\dot{u} : X \to \mathbb{R}$ with $\dot{u}(\alpha) = \dot{u}(\beta) = 0$, let*

$$L(u, \dot{u}) := \int_X l \left((\dot{u} \circ u^{-1})(x), \frac{d}{dx}(\dot{u} \circ u^{-1})(x), \ldots, \frac{d^k}{dx^k}(\dot{u} \circ u^{-1})(x) \right) dx \; .$$

Prove that

- *For any diffeomorphism $\phi : X \to X$, $L(u \circ \phi^{-1}, \dot{u} \circ \phi^{-1})) = L(u, \dot{u})$. This means that L is right-invariant under the action of $Diff^\infty(X)$.*

- *The curve of diffeomorphisms $(t,x) \mapsto u(t,x), [a,b] \times X \to X$ is stationary for the action $\int_a^b L\left(u(t), \frac{du(t)}{dt}\right) dt$ with respect to variations with fixed endpoints if and only if the curve of vector fields $\lambda : [a,b] \times X \to \mathbb{R}$ defined by $\lambda(t, u(t,x)) = \frac{du(t,x)}{dt}$ (i.e. $\lambda := \frac{du}{dt} \circ u^{-1}$) is stationary for the action $\int_X l(\lambda(x), \frac{d\lambda(x)}{dx}, \ldots, \frac{d^k \lambda(x)}{dx^k}) dx$ with respect to variations of the form*

$$\lambda(t,x) + \varepsilon \left(\frac{\partial \delta(t,x)}{\partial t} + \frac{\partial \lambda(t,x)}{\partial x} \delta(t,x) - \frac{\partial \delta(t,x)}{\partial x} \lambda(t,x) \right) \; ,$$

  where $\delta : [a,b] \times X \to \mathbb{R}$ is an arbitrary C^∞ curve of vector fields on X with $\delta(t,\alpha) = \delta(t,\beta) = \delta(a,x) = \delta(b,x) = 0$.

Exercise 4.7 (Burgers' equation as EPDiff) *In exercise 4.6, choose*

$$L(u,\dot{u}) = ||\dot{u}\circ u^{-1}||^2_{L_2} := \int_X \frac{1}{2}(\dot{u}(u^{-1}(x))^2 dx \ .$$

Show that the Euler-Poincaré equations for $\lambda := \frac{\partial u}{\partial t}\circ u^{-1}$ *read*

$$\frac{\partial \lambda}{\partial t} = 3\lambda \frac{\partial \lambda}{\partial x} \ .$$

Exercise 4.8 (The Camassa-Holm equation) *In exercise 4.6, choose*

$$L(u,\dot{u}) = ||\dot{u}\circ u^{-1}||^2_{H^1} = \int_X \frac{1}{2}\left((\dot{u}\circ u^{-1})(x)\right)^2 + \frac{1}{2}\left(\frac{d}{dx}(\dot{u}\circ u^{-1})(x)\right)^2 \ dx \ .$$

Show that this Lagrangian gives rise to the Euler-Poincaré equation

$$\left(1 - \frac{\partial^2}{\partial x^2}\right)\frac{\partial \lambda}{\partial t} = 3\lambda\frac{\partial \lambda}{\partial x} - 2\frac{\partial \lambda}{\partial x}\frac{\partial^2 \lambda}{\partial x^2} - \lambda \frac{\partial^3 \lambda}{\partial x^3} \ .$$

This equation is called the Camassa-Holm equation. It is famous because it is integrable and because it admits a special type of peaked soliton solutions, called 'peakons'.

5 Hamiltonian systems

In this section, we will show that the Euler-Lagrange equations (2.8) for a nondegenerate Lagrangian are equivalent to the famous equations of Hamilton.

5.1 The Legendre transform

Let $L : TQ \to \mathbb{R}$ be such a nondegenerate Lagrangian. We define the "momentum" variables

$$p_j = \frac{\partial L(q, \dot{q})}{\partial \dot{q}_j} \in \mathbb{R} \ .$$

Again under the assumption that $\frac{\partial^2 L(q,\dot{q})}{\partial \dot{q}_j \partial \dot{q}_k}$ is invertible, the transformation

$$(q, \dot{q}) \mapsto (q, p) = (q, \frac{\partial L(q, \dot{q})}{\partial \dot{q}})$$

is a local diffeomorphism. Hence we may write $p_j = p_j(q, \dot{q})$ and $\dot{q}_j = \dot{q}_j(q, p)$. If we now also express the constant of motion (2.9) in terms of the new variables

$$H(q, p) := h(q, \dot{q}(q, p)) = \sum_{j=1}^{n} p_j \dot{q}_j(q, p) - L(q, \dot{q}(q, p)) \ ,$$

then we observe that

$$\frac{\partial H(q, p)}{\partial p_j} = \dot{q}_j(q, p) + \sum_{k=1}^{n} p_k \frac{\partial \dot{q}_k(q, p)}{\partial p_j} - \sum_{k=1}^{n} \frac{\partial L(q, \dot{q}(q, p))}{\partial \dot{q}_k} \frac{\partial \dot{q}_k(q, p)}{\partial p_j} = \dot{q}_j(q, p) \ ,$$

because of the definition of p_k. A similar computation leads to the conclusion that

$$\frac{\partial H(q, p)}{\partial q_j} = -\frac{\partial L(q, \dot{q}(q, p))}{\partial q_j} \ .$$

This makes the Euler-Lagrange equations $\frac{dq_j}{dt} = \dot{q}_j, \frac{d}{dt}\frac{\partial L(q,\dot{q})}{\partial \dot{q}_j} = \frac{\partial L(q,\dot{q})}{\partial q_j}$ for the curve $t \mapsto (q(t), \dot{q}(t))$ in TQ equivalent to the equations

$$\frac{dq_j}{dt} = \frac{\partial H(q, p)}{\partial p_j} \ , \ \frac{dp_j}{dt} = -\frac{\partial H(q, p)}{\partial q_j} \tag{5.1}$$

for the curve $t \mapsto (q(t), p(t))$. Equations (5.1) are called Hamilton's equations of motion for the *Hamiltonian function* H. The transformation of the Lagrangian L into the Hamiltonian H is traditionally called the "Legendre transformation".

The momentum $p(q, \dot{q}) := \frac{\partial L(q,\dot{q})}{\partial \dot{q}}$ does not have the interpretation of an element of $T_q Q$. In fact, $p(q, \dot{q})$ is the total derivative of the function $\dot{q} \mapsto L(q, \dot{q})$ at the point $\dot{q}$. Sometimes it is also called the "fiber derivative" of L as the differentiation is only in the direction of the fiber $T_qQ \subset TQ$. The derivative acts on $v \in T_qQ$ by $p(q,\dot{q})(v) = \frac{d}{d\varepsilon}\big|_{\varepsilon=0} L(q, \dot{q} + \varepsilon v)$, so $p(q,\dot{q})$ is actually a linear map from T_qQ to $\mathbb{R}$, i.e. $p(q,\dot{q}) \in (T_qQ)^*$. Recall that the manifold of all q's and p's is called the cotangent bundle T^*Q of Q: $T^*Q := \cup_{q\in Q}(T_qQ)^*$.

For an arbitrary C^1 function $H : T^*Q \to \mathbb{R}$ of positions and momenta, Hamilton's equations of motion define a vector field on T^*Q, namely

$$X_H := \sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j}\frac{\partial}{\partial p_j} \ . \tag{5.2}$$

X_H is called the "Hamiltonian vector field" of H.

Further reading

Most of the material in these lecture notes can be found in the existing literature in one form or the other, although most of the time with much less detail. For further reading on the subject, I would like to refer you to the following texts.

The bible of Geometric Mechanics, which treats classical mechanics in a differential geometric framework, is

- Abraham, R., Marsden, J.E., *Foundations of Mechanics*. The Benjamin/Cummings Publ. Co., Reading, Mass., 1987.

Easier to read, and with a focus on computations-by-hand is

- Arnol'd, V.I., *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics 60, Springer-Verlag, 1978.

A nice introduction to the application of geometry and topology in fluid mechanics can be found in

- Arnol'd, V.I. and Khesin, B.A., *Topological Methods in Hydrodynamics*, Graduate Texts in Mathematics 60, Springer-Verlag, 1978.

Several examples of classical mechanical systems, treated in a slightly formal way, arise in the textbook

- Cushman, R.H. and Bates, L.M., *Global aspects of Classical Mechanical Systems*, Birkhäuser Verlag, 1997.

An already standard, advanced, but nice book on Lie groups is:

- Duistermaat, J.J. and Kolk, J.A.C., *Lie groups*, Springer-Verlag, 2000.

Finally, introductory texts about the restricted three-body problem are found in

- Meyer, K.R., *Periodic Solutions of the N-Body-Problem*, Lecture Notes in Mathematics, Springer-Verlag, 1999.

- Meyer, K.R., Hall, G.R., *Introduction to Hamiltonian dynamical systems and the N-body problem*, Applied Math. Sciences 90, Springer-Verlag, 1992.