Mathematical Aspects of Classical Mechanics

Yannis Bähni*

*ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland. *E-mail address*: baehniy@student.ethz.ch

Preface

This notes are the product of a semester project done at the ETH Zurich in the autumn semester of 2018 under the supervision of *Dr. Ana Cannas da Silva*[†]. I will roughly follow the first chapter of the book *Quantum Mechanics for Mathematicians* by *Leon A. Takhtajan* [Tak08], which serves as a brief introduction to classical mechanics. Since this introduction is very brief, understandable by considering its purpose, I additionally rely on the classic *Mathematical Methods of Classical Mechanics* by *Vladimir I. Arnold* [Arn89]. As the title already suggests, this is not a treatment of the physical part of classical mechanics, but a mathematical one. Hence the aim of these notes is to give a thoughtful introduction to the mathematical methods used in the realm of classical mechanics and their strong connection to differential topology and differential geometry, especially *symplectic geometry*.

Winterthur, September 8, 2018 — Yannis Bähni

[†]Departement of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland.

Contents

CHAPTER 1

Lagrangian Mechanics

Introduction

Classical mechanics deals with differential equations originating from extremals of *functionals*, i.e. functions defined on an infinite-dimensional function space. The study of such extremality properties of functionals is known as the *calculus of variations*. To illustrate this fundamental principle, let us consider the *variational formulation* of second order elliptic operators in divergence form based on [Str14, pp. 167–168].
Let $n \in \mathbb{N}$, $n \geq 1$, and $\Omega \subseteq\subseteq \mathbb{R}^n$ such that $\bar{\Omega}$ is a manifold with boundary. Moreover, let $H_0^1(\Omega)$ denote the Sobolev space $W_0^{1,2}(\Omega)$ with inner product

$$\langle u, v\rangle_{H_0^1(\Omega)} = \int_\Omega uv + \int_\Omega \nabla u \nabla v.$$

Suppose $a^{ij} \in C^\infty(\bar{\Omega})$ symmetric, $f \in C^\infty(\bar{\Omega})$ and consider the second order homogenous Dirichlet problem

$$\begin{cases} -\dfrac{\partial}{\partial x^j}\left(a^{ij}\dfrac{\partial u}{\partial x^i}\right) = f & \text{in } \Omega\,, \\ u = 0 & \text{on } \partial\Omega\,, \end{cases} \tag{1}$$

Suppose $u \in C^\infty(\bar{\Omega})$ solves (1). Then integration by parts (see [Lee13, p. 436]) yields

$$\int_\Omega fv = -\int_\Omega \frac{\partial}{\partial x^j}\left(a^{ij}\frac{\partial u}{\partial x^i}\right)v = -\int_\Omega \operatorname{div}(X)v = \int_\Omega \langle X, \nabla v\rangle = \int_\Omega a^{ij}\frac{\partial u}{\partial x^i}\frac{\partial v}{\partial x^j}$$

for any $v \in C_c^\infty(\Omega)$, where $X := \left(a^{ij}\frac{\partial u}{\partial x^i}\right)_j$. Thus we say that $u \in H_0^1(\Omega)$ is a *weak solution* of (1) iff

$$\forall v \in C_c^\infty(\Omega) : \int_\Omega a^{ij}\frac{\partial u}{\partial x^i}\frac{\partial v}{\partial x^j} = \int_\Omega fv.$$

If $(a^{ij})_{ij}$ is *uniformly elliptic*, i.e. there exists $\lambda > 0$ such that

$$\forall x \in \Omega \forall \xi \in \mathbb{R}^n : a^{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2\,,$$

then (1) admits a unique weak solution $u \in H_0^1(\Omega)$ (in fact $u \in C^\infty(\Omega)$ using *regularity theory*, for more details see [Str14, p. 175]). Indeed, observe that

$$\langle\cdot,\cdot\rangle_a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$$

defined by

$$\langle u, v\rangle_a := \int_\Omega a^{ij} \frac{\partial u}{\partial x^i} \frac{\partial v}{\partial x^j} \tag{2}$$

is an inner product on $H_0^1(\Omega)$ with induced norm equivalent to the standard one on $H_0^1(\Omega)$ due to Poincaré's inequality [Str14, p. 107]. Applying the Riesz Representation theorem [Str14, pp. 49–50] yields the result. Moreover, this solution can be characterized by a *variational principle*, i.e. if we define the *energy functional* $E : H_0^1(\Omega) \to \mathbb{R}$

$$E(v) := \frac{1}{2}\|v\|_a^2 - \int_\Omega f v,$$

for any $v \in H_0^1(\Omega)$, where $\|\cdot\|_a$ denotes the norm induced by the inner product (2), then $u \in H_0^1(\Omega)$ solves (1) if and only if

$$E(u) = \inf_{v \in H_0^1(\Omega)} E(v). \tag{3}$$

Indeed, suppose $u \in H_0^1(\Omega)$ is a solution of (1). Let $v \in H_0^1(\Omega)$. Then $u = v + w$ for $w := u - v \in H_0^1(\Omega)$ and we compute

$$E(v) = E(u+w) = \frac{1}{2}\|u\|_a^2 + \langle u, w\rangle_a + \frac{1}{2}\|w\|_a^2 - \int_\Omega f(u+w) = E(u) + \frac{1}{2}\|w\|_a^2 \geq E(u)$$

with equality if and only if $u = v$ a.e. Conversly, suppose the infimum is attained by some $u \in H_0^1(\Omega)$. Thus by elementary calculus

$$0 = \left.\frac{d}{dt}\right|_{t=0} E(u + tv) = \langle u, v\rangle_a - \int_\Omega f v \tag{4}$$

for all $v \in H_0^1(\Omega)$.

Suppose now that $u \in C^\infty(\overline{\Omega})$ with $u|_{\partial\Omega} = 0$ solves the variational formulation (3). Then again integration by parts yields

$$\langle u, v\rangle_a - \int_\Omega f v = -\int_\Omega \operatorname{div}(X) v - \int_\Omega f v = \int_\Omega \left(-\frac{\partial}{\partial x^j}\left(a^{ij}\frac{\partial u}{\partial x^i}\right) - f\right) v$$

for all $v \in C_c^\infty(\Omega)$ and where $X := \left(a^{ij}\frac{\partial u}{\partial x^i}\right)_j$. Hence (4) implies

$$\forall v \in C_c^\infty(\Omega) : \int_\Omega \left(-\frac{\partial}{\partial x^j}\left(a^{ij}\frac{\partial u}{\partial x^i}\right) - f\right) v = 0.$$

We might expect that this implies

$$-\frac{\partial}{\partial x^j}\left(a^{ij}\frac{\partial u}{\partial x^i}\right) = f.$$

That this is indeed the case, is guaranteed by a foundational result in the *calculus of variations* (therefore the name).

Proposition 1.1 (Fundamental Lemma of Calculus of Variations). *Let* $\Omega \subseteq \mathbb{R}^n$ *open and* $f \in L^1_{\mathrm{loc}}(\Omega)$*. If*

$$\forall \varphi \in C_c^\infty(\Omega) : \int_\Omega f\varphi = 0,$$

then $f = 0$ *a.e.*

Proof. See [Str14, p. 40]. □

Thus we recovered a second order partial differential equation from the variational formulation. In fact, this is exactly the boundary value problem (1) from the beginning of our exposition. This technique, and in particular the fundamental lemma of calculus of variations 1.1 will play an important role in our treatment of classical mechanics. However, since we are concerned with smooth manifolds only, we use a version of the fundamental lemma of calculus of variations 1.1, which is fairly easy to prove and hence really deserves the terminology "lemma".

Lemma 1.2 (Fundamental Lemma of Calculus of Variations, Smooth Version). *Let* $\Omega \subseteq \mathbb{R}^n$ *open and* $f \in C^\infty(\Omega)$*. If*

$$\forall \varphi \in C_c^\infty(\Omega) : \int_\Omega f\varphi = 0,$$

then $f = 0$*.*

Proof. Towards a contradiction, assume that $f \neq 0$ on Ω. Thus there exists $x_0 \in \Omega$, such that $f(x_0) \neq 0$. Without loss of generality, we may assume that $f(x_0) > 0$, since otherwise, consider $-f$ instead of f. The smoothness of f implies the continuity of f on Ω. Thus there exists $\delta > 0$, such that $f(x) \in B_{f(x_0)/2}\left(f(x_0)\right)$ holds for all $x \in B_\delta(x_0)$ or equivalently, $f(x) > f(x_0)/2 > 0$ for all $x \in B_\delta(x_0)$. By lemma 2.22 [Lee13, p. 42], there exists a smooth bump function φ supported in $B_\delta(x_0)$ and $\varphi = 1$ on $\overline{B}_{\delta/2}(x_0)$. In particular, $\varphi \in C_c^\infty(\Omega)$. Therefore we have

$$0 = \int_\Omega f\varphi = \int_{B_\delta(x_0)} f\varphi \geq \int_{B_{\delta/2}(x_0)} f\varphi > \frac{1}{2} f(x_0)\left|B_{\delta/2}(x_0)\right| > 0,$$

which is a contradiction. □

Exercise 1.3. [1] Let $\Omega \subseteq\subseteq \mathbb{R}^n, 2 \leq p < \infty$ and define $\mathcal{B} := \{v \in C^\infty(\overline{\Omega}) : v|_{\partial\Omega} = 0\}$. Moreover, define $E_p : \mathcal{B} \to \mathbb{R}$ by $E_p(v) := \int_\Omega |\nabla v|^p$. Derive the partial differential equation satisfied by minimizers $u \in \mathcal{B}$ of the variational problem $E(u) = \inf_{v \in \mathcal{B}} E(v)$.

[1] This is exercise 1.2.(*b*) from exercise sheet 1 of the course *Functional Analysis II* taught by *Prof. Dr. A. Carlotto* at ETHZ in the spring of 2018, which can be found here.

Lagrangian Systems and the Principle of Least Action

Mechanical systems, i.e. a pendulum, are modelled using the language of differential geometry. Thus it is necessary to introduce the relevant physical counterparts.

Definition 1.4 (Configuration Space). *A **configuration space** is defined to be a finite-dimensional manifold in* Diff.

Definition 1.5 (Motion). *A **motion in a configuration space M** is defined to be a path $\gamma \in C^\infty(J, M)$, where $J \subseteq \mathbb{R}$ is an interval.*

Definition 1.6 (State). *A **state of the configuration space** is defined to be an element of the tangent bundle of the configuration space, called the **state space**.*

Also in classical mechanics, one has to rely on basic principles, which are to some extent experimentally verified. One of the most fundamental is the following.

Axiom 1 (Newton-Laplace Determinacy Principle). *A motion in a configuration space is completely determined by a state at some instant.*

The Newton-Laplace determinacy principle 1 motivates our main definition of this chapter.

Definition 1.7 (Lagrangian System). *A **Lagrangian system** is defined to be a tuple (M, L) consisting of an object $M \in$ Diff and a function $L \in C^\infty(TM \times \mathbb{R})$, called a **Lagrangian function**.*

Example 1.8. For $M \in$ Diff let $T \in C^\infty(TM \times \mathbb{R})$ and $V \in C^\infty(M \times \mathbb{R})$. Define $L \in C^\infty(TM \times \mathbb{R})$ by $L := T - V$. In this situation, T is called the ***kinetic energy*** and V is called the ***potential energy***.

Definition 1.9 (Path Space). *Let $M \in$ Diff, $q_0, q_1 \in M$ and $t_0, t_1 \in \mathbb{R}$ with $t_0 \leq t_1$. Define the **path space of M connecting (q_0, t_0) and (q_1, t_1)** to be the set*

$$\mathcal{P}(M)^{q_0,t_0}_{q_1,t_1} := \{\gamma \in C^\infty([t_0, t_1], M) : \gamma(t_0) = q_0 \text{ and } \gamma(t_1) = q_1\}. \tag{5}$$

Remark 1.10. For the sake of simplicity, we will just use the terminology *path space* for $\mathcal{P}(M)^{q_0,t_0}_{q_1,t_1}$ and simply write $\mathcal{P}(M)$. We implicitely assume the conditions of definition 1.9, however.

Definition 1.11 (Variation). *Let $\mathcal{P}(M)$ be a path space and $\gamma \in \mathcal{P}(M)$. A **variation of γ** is defined to be a morphism $\Gamma \in C^\infty([t_0, t_1] \times [-\varepsilon_0, \varepsilon_0], M)$ for some $\varepsilon_0 > 0$ and such that*

- *$\Gamma(t, 0) = \gamma$ for all $t \in [t_1, t_0]$.*
- *$\Gamma(t_0, \varepsilon) = q_0$ for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.*
- *$\Gamma(t_1, \varepsilon) = q_1$ for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.*

Remark 1.12. If Γ is a variation of $\gamma \in \mathcal{P}(M)$, we write $\gamma_\varepsilon(-) := \Gamma(-, \varepsilon)$ for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.

Example 1.13 (Perturbation of a Path along a Single Direction). Let $M \in \mathsf{Diff}$ of dimension n, (U, φ) a chart and suppose that γ is a path in U. With respect to this chart, we can write the coordinate representation of γ as

$$\gamma(t) = \left(\gamma^1(t), \dots, \gamma^n(t)\right)$$

for any $t \in [t_0, t_1]$. Let $f \in C_c^\infty(t_0, t_1)$. Consider the family $\Gamma : [t_0, t_1] \times [-\varepsilon_0, \varepsilon_0] \to M$ defined by

$$\Gamma(t, \varepsilon) := (\iota \circ \varphi^{-1}) \left(\gamma^1(t), \dots, \gamma^i(t) + \varepsilon f(t), \dots, \gamma^n(t)\right)$$

where $\iota : U \hookrightarrow M$ denotes inclusion and $\varepsilon_0 > 0$ is to be determined. By exercise 1.14, there exists $\delta > 0$ such that

$$U_\delta := \{x \in \mathbb{R}^n : \operatorname{dist}(x, \gamma([t_0, t_1])) < \delta\} \subseteq \varphi(U).$$

Choose $\varepsilon_0 > 0$ such that $0 < \varepsilon_0 < \delta / \|f\|_\infty$. Then in coordinates

$$\operatorname{dist}\left(\gamma_\varepsilon(t), \gamma([t_0, t_1])\right) \leq |\gamma_\varepsilon(t) - \gamma(t)| = |\varepsilon| \|f\|_\infty \leq \varepsilon_0 \|f\|_\infty < \delta$$

for all $t \in [t_0, t_1]$. Hence $\gamma_\varepsilon(t) \in U_\delta$ and thus $\gamma_\varepsilon(t) \in \varphi(U)$. Therefore, Γ is indeed well-defined. Moreover, it is easy to show that the properties of definition 1.11 holds, therefore, Γ is a variation of γ. In fact, this example shows, that any path γ contained in a single chart admits infinitely many variations. An example of such a variation is shown in figure 1.

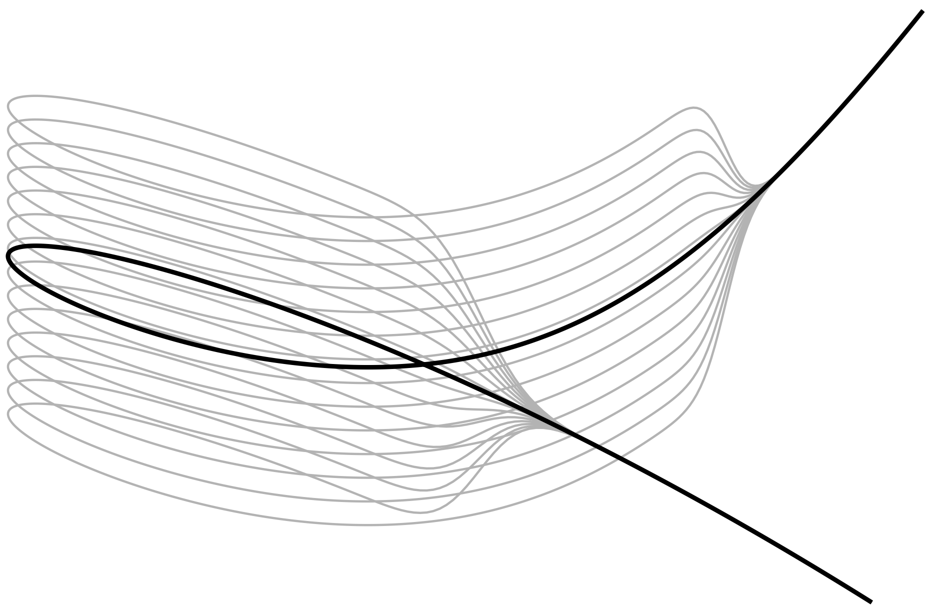

Figure 1. Example of a variation of the path $\gamma(t) = (\gamma^1(t), \gamma^2(t))$ in $\mathbb{R}^2$ defined by $\gamma(t) := (t^2 + \sin(t)\cos(t), t^3 - t)$ for $t \in [-\frac{3}{2}, \frac{3}{2}]$ along the second coordinate using a smooth bump function as in [Lee13, p. 42].

Exercise 1.14. Let $U \subseteq \mathbb{R}^n$ open and $A \subseteq U$ closed. Then there exists $\delta > 0$ such that

$$U_\delta := \{x \in \mathbb{R}^n : \mathrm{dist}(x, A) < \delta\} \subseteq U.$$

Definition 1.15 (Action Functional). *Let (M, L) be a Lagrangian system and $\mathcal{P}(M)$ be a path space. The morphism $S : \mathcal{P}(M) \to \mathbb{R}$ defined by*

$$S(\gamma) := \int_{t_0}^{t_1} L\left(\gamma(t), \dot{\gamma}(t), t\right) dt$$

is called the ***action functional****.*

Motions of Lagrangian systems are characterized by an axiom.

Axiom 2 (Hamilton's Principle of Least Action). *Let (M, L) be a Lagrangian system and $\mathcal{P}(M)$ be a path space. A path $\gamma \in C^\infty([t_0, t_1], M)$ describes a motion of (M, L) between (q_0, t_0) and (q_1, t_1) if and only if*

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\gamma_\varepsilon) = 0 \tag{6}$$

for all variations γ_ε of γ.

Definition 1.16 (Extremal). *A motion of a Lagrangian system between two points is called an* ***extremal of the action functional S****.*

The Newton-Laplace determinacy principle 1 implies that motions of mechanical systems can be described as solutions of second order ordinary differential equations. That this is indeed the case, is shown by the next theorem.

Theorem 1.17 (Euler-Lagrange Equations). *Let (M, L) be a Lagrangian system. If a path $\gamma \in C^\infty([t_0, t_1], M)$ describes a motion of (M, L) between (q_0, t_0) and (q_1, t_1) then for all charts (U, x^i)*

$$\frac{\partial L}{\partial x}\big(\gamma(t), \dot{\gamma}(t), t\big) - \frac{d}{dt}\frac{\partial L}{\partial v}\big(\gamma(t), \dot{\gamma}(t), t\big) = 0 \tag{7}$$

holds, where (x^i, v^i) denotes the standard coordinates on TM. The system of equations (7) is referred to as the ***Euler-Lagrange equations****.*

Proof. By Hamilton's principle of least action 2, we may assume that γ is an extremal of the action functional S. The proof is divided into two steps.

Step 1: Suppose that γ of S is conatined in a chart domain U. Let $t \in [t_0, t_1]$ and abreviate $p := (\gamma(t), \dot{\gamma}(t), t)$ Using the formula for the derivative of a function along a curve [Lee13, p. 283], we compute

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} L\left(\gamma_\varepsilon(t), \dot{\gamma}_\varepsilon(t), t\right) = dL_p\left(\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \gamma_\varepsilon(t), \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \dot{\gamma}_\varepsilon(t), 0\right)$$

$$= dL_p \left(\frac{d\gamma_\varepsilon^j(t)}{d\varepsilon}(0) \frac{\partial}{\partial x^j}\bigg|_{\gamma(t)}, \frac{d\dot{\gamma}_\varepsilon^j(t)}{d\varepsilon}(0) \frac{\partial}{\partial v^j}\bigg|_{\dot{\gamma}(t)}, 0 \right).$$

for all variations γ_ε of γ in U. Moreover

$$dL_p = \frac{\partial L}{\partial x^i}(p) dx^i|_p + \frac{\partial L}{\partial v^i}(p) dv^i|_p + \frac{\partial L}{\partial t}(p) dt|_p.$$

Therefore

$$\begin{aligned}
0 &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} S(\gamma_\varepsilon) \\
&= \int_{t_0}^{t_1} \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} L\left(\gamma_\varepsilon(t), \dot{\gamma}_\varepsilon(t), t\right) dt \\
&= \int_{t_0}^{t_1} dL_p \left(\frac{d\gamma_\varepsilon^j(t)}{d\varepsilon}(0) \frac{\partial}{\partial x^j}\bigg|_{\gamma(t)}, \frac{d\dot{\gamma}_\varepsilon^j(t)}{d\varepsilon}(0) \frac{\partial}{\partial v^j}\bigg|_{\dot{\gamma}(t)}, 0 \right) \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial x^i}(p) \frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0) dt + \int_{t_0}^{t_1} \frac{\partial L}{\partial v^i}(p) \frac{d\dot{\gamma}_\varepsilon^i(t)}{d\varepsilon}(0) dt \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial x^i}(p) \frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0) dt + \int_{t_0}^{t_1} \frac{\partial L}{\partial v^i}(p) \left(\frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0) \right)' dt \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial x^i}(p) \frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0) dt + \frac{\partial L}{\partial v^i}(p) \frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0)\bigg|_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt} \frac{\partial L}{\partial v^i}(p) \frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0) dt \\
&= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x^i}(p) - \frac{d}{dt} \frac{\partial L}{\partial v^i}(p) \right) \frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0) dt
\end{aligned}$$

since $\gamma_\varepsilon^i(t_0)$ and $\gamma_\varepsilon^i(t_1)$ are constant by definition of a variation. Let $f \in C_c^\infty(t_0, t_1)$, $j = 1, \ldots, n$ and γ_ε be the variation of γ defined in example 1.13 along the j-th direction. Above computation therefore yields

$$0 = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x^j}(p) - \frac{d}{dt} \frac{\partial L}{\partial v^j}(p) \right) f(t) dt$$

for all $f \in C_c^\infty(t_0, t_1)$. Hence the fundamental lemma of calculus of variations 1.2 implies

$$\frac{\partial L}{\partial x^j}(p) - \frac{d}{dt} \frac{\partial L}{\partial v^j}(p) = 0$$

for all $j = 1, \ldots, n$.

Step 2: Suppose that γ is an arbitrary extremal of S. The key technical result used here is the following lemma.

Lemma 1.18 (Lebesgue Number Lemma). *Every open cover of a compact metric space admits a Lebesgue number, i.e. a number $\delta > 0$ such that every subset of the metric space with diameter less than δ is contained in a member of the family.*

Proof. See [Lee11, p. 194]. □

Let $(U_\alpha)_{\alpha\in A}$ be the smooth structure on M, i.e. the maximal smooth atlas. Since γ is continuous, $\left(\gamma^{-1}(U_\alpha)\right)_{\alpha\in A}$ is an open cover for $[t_0, t_1]$. By the Lebesgue number lemma 1.18, this open cover admits a Lebesgue number $\delta > 0$. Let $k \in \mathbb{N}$ such that $(t_1 - t_0)/k < \delta$ and define

$$t_i := \frac{i}{k}(t_1 - t_0) + t_0$$

for all $i = 0, \ldots, k$. Then for all $i = 1, \ldots, k$, $\gamma|_{[x_{i-1}, x_i]}$ is contained in U_α for some $\alpha \in A$. Hence applying step 1 yields the result.

□

Remark 1.19. If a configuration space can be covered by a single chart, then the statement of theorem 1.17 becomes an equivalence.

Due to the Newton-Laplace Determinacy Principle 1, the motions on a Lagrangian system are inherently characterized by the Lagrangian function and locally by the Euler-Lagrange equations (7). Hence any motion satisfies locally a system of second order ordinary differential equations. This system bears its own name.

Definition 1.20 (Equations of Motion). *The Euler-Lagrange equations* (7) *of a Lagrangian system are called the* ***equations of motion****.*

Example 1.21 (Motions on Riemannian Manifolds). Let (M, g) be a Riemannian manifold of dimension n and consider the Lagrangian L on M defined in example 1.8 with kinetic energy

$$T(q, v, t) := \frac{1}{2} g_q(v, v) = \frac{1}{2}|v|_g^2$$

and potential energy $V(q, t) := 0$ for $q \in M$, $v \in T_q M$ and $t \in \mathbb{R}$. Let (U, x^i) be a chart on M. We compute

$$\begin{aligned} L(q, v, t) &= \frac{1}{2} g_q\,(v, v) \\ &= \frac{1}{2} g_q \left(v^i \left.\frac{\partial}{\partial x^i}\right|_q, v^j(t) \left.\frac{\partial}{\partial x^j}\right|_q \right) \\ &= \frac{1}{2} g_q \left(\left.\frac{\partial}{\partial x^i}\right|_q, \left.\frac{\partial}{\partial x^j}\right|_q \right) v^i v^j \\ &= \frac{1}{2} g_{ij}(q) v^i v^j, \end{aligned}$$

where $g_{ij}(q) := g_q\left(\left.\frac{\partial}{\partial x^i}\right|_q, \left.\frac{\partial}{\partial x^j}\right|_q \right)$. Thus $\frac{\partial L}{\partial x^l}(q, v, t) = \frac{1}{2}\frac{\partial g_{ij}}{\partial x^l}(q) v^i v^j$ and in particular

$$\frac{\partial L}{\partial x^l}\left(\gamma(t), \dot\gamma(t), t\right) = \frac{1}{2}\frac{\partial g_{ij}}{\partial x^l}\left(\gamma(t)\right)\dot\gamma^i(t)\dot\gamma^j(t),$$

for all $l = 1, \dots, n$. Moreover

$$\frac{\partial L}{\partial v^l}(q,v,t) = \frac{1}{2} g_{ij}(q)\delta^i_l v^j + \frac{1}{2} g_{ij}(q) v^i \delta^j_l = \frac{1}{2} g_{lj}(q) v^j + \frac{1}{2} g_{il}(q) v^i$$

implies

$$\begin{aligned}
\frac{d}{dt}\frac{\partial L}{\partial v^l}\big(\gamma(t),\dot\gamma(t),t\big) &= \frac{1}{2}\frac{d}{dt} g_{lj}(\gamma)\dot\gamma^j + \frac{1}{2} g_{lj}(\gamma)\ddot\gamma^j + \frac{1}{2}\frac{d}{dt} g_{il}(\gamma)\dot\gamma^i + \frac{1}{2} g_{il}(\gamma)\ddot\gamma^i \\
&= \frac{1}{2} dg_{lj}(\dot\gamma)\dot\gamma^j + \frac{1}{2} g_{lj}(\gamma)\ddot\gamma^j + \frac{1}{2} dg_{il}(\dot\gamma)\dot\gamma^i + \frac{1}{2} g_{il}(\gamma)\ddot\gamma^i \\
&= \frac{1}{2}\frac{\partial g_{lj}}{\partial x^k}\dot\gamma^k\dot\gamma^j + \frac{1}{2} g_{lj}(\gamma)\ddot\gamma^j + \frac{1}{2}\frac{\partial g_{il}}{\partial x^k}\dot\gamma^k\dot\gamma^i + \frac{1}{2} g_{il}(\gamma)\ddot\gamma^i \\
&= \frac{1}{2}\frac{\partial g_{jl}}{\partial x^k}\dot\gamma^k\dot\gamma^j + \frac{1}{2} g_{jl}(\gamma)\ddot\gamma^j + \frac{1}{2}\frac{\partial g_{il}}{\partial x^k}\dot\gamma^k\dot\gamma^i + \frac{1}{2} g_{il}(\gamma)\ddot\gamma^i \\
&= g_{il}\ddot\gamma^i + \frac{1}{2}\frac{\partial g_{jl}}{\partial x^i}\dot\gamma^i\dot\gamma^j + \frac{1}{2}\frac{\partial g_{il}}{\partial x^j}\dot\gamma^i\dot\gamma^j.
\end{aligned}$$

Therefore the Euler-Lagrange equations (7) read

$$0 = \frac{d}{dt}\frac{\partial L}{\partial v^l} - \frac{\partial L}{\partial x^l} = g_{il}\ddot\gamma^i + \frac{1}{2}\left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l}\right)\dot\gamma^i\dot\gamma^j,$$

for all $l = 1, \dots, n$. Multiplying both sides by g^{kl} yields

$$\ddot\gamma^k + \Gamma^k_{ij}\dot\gamma^i\dot\gamma^j = 0, \tag{8}$$

for all $k = 1, \dots, n$, where

$$\Gamma^k_{ij} := \frac{1}{2} g^{kl}\left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l}\right)$$

are the ***Christoffel symbols*** with respect to the choosen chart (see [Lee97, p. 70]). Equation (8) is called the ***geodesic equation*** (see [Lee97, p. 58]). Hence extremals γ of the action functional satisfy the geodesic equation and are therefore geodesics on the Riemannian manifold M.

Lemma 1.22. *Let (M, L) be a Lagrangian system and define $L + df \in C^\infty(TM \times \mathbb{R})$ by*

$$(L + df)(q,v,t) := L(q,v,t) + df_q(v)$$

for any $f \in C^\infty(M)$. Then (M, L) and $(M, L + df)$ admit the same equations of motion.

Proof. Let us denote the action function corresponding to $L + df$ by $\widetilde{S}$ and suppose γ_ε is a variation of γ in M. Using the formula for the derivative of a function along a curve [Lee13, p. 283] we compute

$$\widetilde{S}(\gamma_\varepsilon) = \int_{t_0}^{t_1} L(\gamma_\varepsilon(t),\dot\gamma_\varepsilon(t),t)dt + \int_{t_0}^{t_1} df_{\gamma_\varepsilon(t)}\big(\dot\gamma_\varepsilon(t)\big)\,dt$$

$$
\begin{aligned}
&= S(\gamma_\varepsilon) + \int_{t_0}^{t_1} (f \circ \gamma_\varepsilon)'(t)dt \\
&= S(\gamma_\varepsilon) + f\big(\gamma_\varepsilon(t_1)\big) - f\big(\gamma_\varepsilon(t_0)\big) \\
&= S(\gamma_\varepsilon) + f(q_1) - f(q_0).
\end{aligned}
$$

In particular

$$
\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \widetilde{S}(\gamma_\varepsilon) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\gamma_\varepsilon).
$$

□

Remark 1.23. Lemma 1.22 implies, that the Lagrangian of a mechanical system can only be determined up to differentials of smooth functions. Actually, in coordinates, also up to total time derivatives. Hence a *law of motion*, that is a Lagrangian describing a certain mechanical system, is in fact an equivalence class of Lagrangian functions.

Legendre Transform

In this section we *dualize* the notion of a Lagrangian function, that is, to each Lagrangian function $L \in C^\infty(TM)$ we will associate a *dual function* $L^* \in C^\infty(T^*M)$. It turns out, that in this dual formulation, the equations of motion take a very symmetric form. To simplify the notation and illuminating the main concept, we consider Lagrangian functions of a special type.

Definition 1.24 (Autonomous System). *A Lagrangian system (M, L) is said to be an* ***autonomous Lagrangian system****, iff $L \in C^\infty(TM)$.*

The Energy of a Lagrangian System. Let $M \in \mathsf{Diff}$ of dimension n and $(U, (x^i))$ a chart on M. Let (x^i, v^i) denote standard coordinates on TM, that is $v^i := dx^i$ for all $i = 1, \dots, n$. Suppose $L \in C^\infty(TM)$. For every $(x, v) \in TM$ we can define a covector $d^{\mathcal{F}}_{(x,v)}L \in T^*_x M$ by setting

$$
d^{\mathcal{F}}_{(x,v)}L := \left.\frac{\partial}{\partial v^i}\right|_{(x,v)} (L) dx^i|_x = \frac{\partial L}{\partial v^i} dx^i. \tag{9}
$$

Let $(\widetilde{x}^i, \widetilde{v}^i)$ denote another pair of coordinates on TM. Then we have that

$$
\frac{\partial}{\partial \widetilde{v}^i} = \frac{\partial x^j}{\partial \widetilde{v}^i}\frac{\partial}{\partial x^j} + \frac{\partial v^j}{\partial \widetilde{v}^i}\frac{\partial}{\partial v^j} = \frac{\partial v^j}{\partial \widetilde{v}^i}\frac{\partial}{\partial v^j}.
$$

Moreover

$$
\frac{\partial}{\partial x^j} = \frac{\partial \widetilde{x}^k}{\partial x^j}\frac{\partial}{\partial \widetilde{x}^k}
$$

which implies

$$
d\widetilde{x}^i\left(\frac{\partial}{\partial x^j}\right) = \frac{\partial \widetilde{x}^k}{\partial x^j} d\widetilde{x}^i\left(\frac{\partial}{\partial \widetilde{x}^k}\right) = \frac{\partial \widetilde{x}^k}{\partial x^j}\delta^i_k = \frac{\partial \widetilde{x}^i}{\partial x^j}.
$$

Thus

$$d\tilde{x}^i = \frac{\partial \tilde{x}^i}{\partial x^j} dx^j$$

or equivalently

$$v^j = \frac{\partial x^j}{\partial \tilde{x}^i} \tilde{v}^i,$$

and so we compute

$$d^{\mathcal{F}} L = \frac{\partial L}{\partial \tilde{v}^i} d\tilde{x}^i = \frac{\partial v^j}{\partial \tilde{v}^i} \frac{\partial L}{\partial v^j} \frac{\partial \tilde{x}^i}{\partial x^k} dx^k = \frac{\partial x^j}{\partial \tilde{x}^i} \frac{\partial L}{\partial v^j} \frac{\partial \tilde{x}^i}{\partial x^k} dx^k = \frac{\partial L}{\partial v^j} \delta^j_k dx^k = \frac{\partial L}{\partial v^j} dx^j.$$

Therefore, $d^{\mathcal{F}}$ is independent of the choice of coordinates.

Definition 1.25 (Fibrewise Differential[2]). *Let (M, L) be an autonomous Lagrangian system. The map $d^{\mathcal{F}} L : TM \to T^*M$ defined on a chart (U, x^i) of M by*

$$d^{\mathcal{F}} L := \frac{\partial L}{\partial v^i} dx^i \tag{10}$$

*where (x^i, v^i) denotes the induced standard coordinates on TM, is called the **fibrewise differential of L**.*

Remark 1.26. The preceeding discussion showed, that the fibrewise differential $d^{\mathcal{F}} L$ is well-defined.

Definition 1.27 (Energy). *The **energy** of an autonomous Lagrangian system (M, L) is defined to be the function $E_L \in C^\infty(TM)$ given by*

$$E_L(x, v) := d^{\mathcal{F}}_{(x,v)} L(v) - L(x, v),$$

in standard coordinates (x^i, v^i) of TM.

Example 1.28. Let (M, g) be a Riemannian manifold and consider the Lagrangian $T - U$, with kinetic energy $T \in C^\infty(TM)$ defined by $T(v) := \frac{1}{2}|v|^2$ and potential energy $U \in C^\infty(M)$. Then the computation performed in example 1.21 yields

$$\begin{aligned} E_{T-U}(x, v) &= \frac{\partial T}{\partial v^k} v^k - \frac{\partial U}{\partial v^k} v^k - T(v) + U(x) \\ &= \frac{1}{2} g_{ij} \delta^i_k v^j v^k + \frac{1}{2} g_{ij} v^i \delta^j_k v^k - T(v) + U(x) \\ &= g_{ij} v^i v^j - T(v) + U(x) \\ &= T(v) + U(x) \end{aligned}$$

for every $(x, v) \in TM$. Hence the energy of this Lagrangian system is given by *kinetic energy plus potential energy*.

Hamilton's Equations.

[2]This terminology is adapted from exercise C.3. on problem sheet C of the lecture *Differential geometry I* taught by *Will J. Merry* at *ETH Zürich* in the autumn semester 2018, which can be found here.

CHAPTER 2

Hamiltonian Mechanics