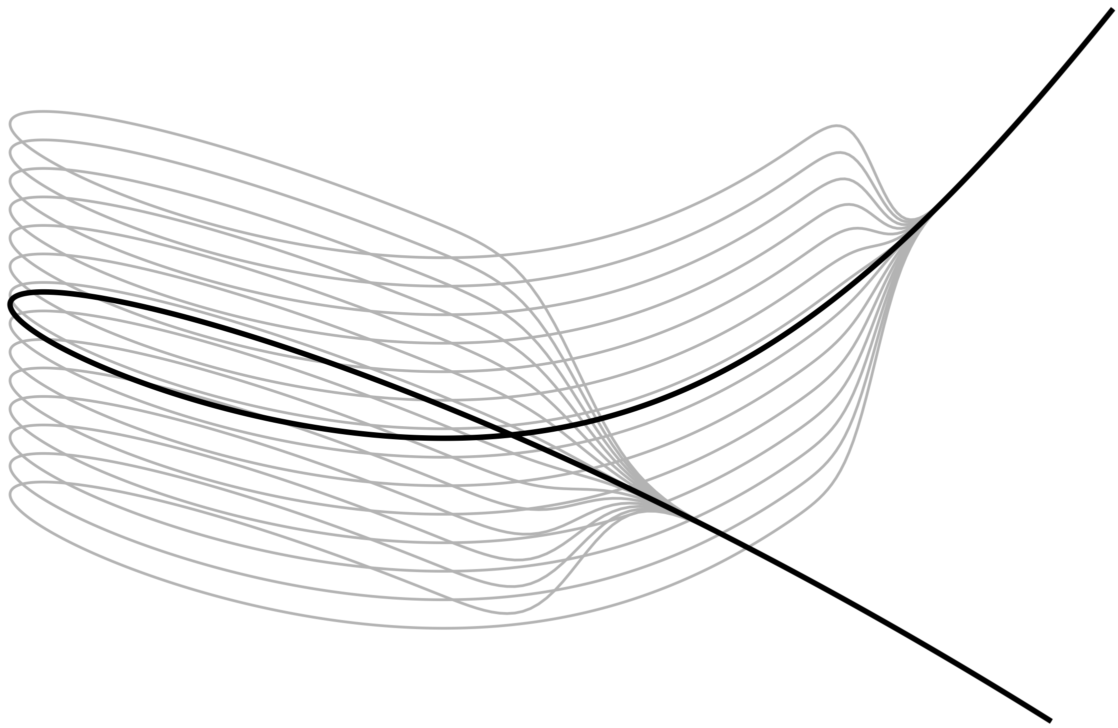
# Mathematical Aspects of Classical Mechanics

Yannis Bähni



ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

# Preface

This notes are the product of a semester project done at the *ETH Zürich* 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 an 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 rather 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*. Therefore it is only natural to consider also the book *Lectures on Symplectic Geometry* by *Ana Cannas da Silva* [Sil08].
I would like to thank first of all my supervisor Dr. Cannas da Silva for granting me this opportunity of writing these notes, and also for introducting me to symplectic geometry back in the autumn semester 2017. Moreover, I would like to thank *Prof. Dr. Will J. Merry*, whose brilliant lectures on *Algebraic Topology* as well as *Differential Geometry* helped me alot in understanding this and related subjects. Also, he was a great help in answering questions and clarifying concepts. A big help was also the marvelous trilogy of books from *John M. Lee* ([Lee11], [Lee13] and [Lee97]), which clear, thoughtful and highly formal exposition of the subject give an in-depth understanding of the matter. I won't deny the obvious: My style of writing and even the typeset of this document is highly inspired, sometimes even copied, from the style used by Jack Lee. The simple reason is, that I appreciate his work very much and try to achieve the same fineness. A prominent indicator of this fact is also the numerous citations of his books in these notes. Lastly, I would like to thank both the mathematics institute at the *University of Zürich* as well as the mathematics institute here at ETH Zürich, for teaching me mathematics. Without whom, maybe I would never have experienced the passion for doing mathematics. In this sense, happy reading (shamelessly ripped off the preface of [Lee13])!

Winterthur, Yannis Bähni
September 8, 2018

# Contents

CHAPTER 1

# Lagrangian Mechanics

## Introduction

Classical mechanics deals with ordinary differential equations originating from extremals of ***functionals***, that is 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].
For convention, unless expliciety stated otherwise, we will assume that all manifolds are smooth, that is of class $C^\infty$, finite-dimensional, Hausdorff and paracompact with at most countably many connected components.
Let $n \in \mathbb{N}$, $n \geq 1$, and $\Omega \subseteq\subseteq \mathbb{R}^n$ such that $\bar{\Omega}$ is a smooth 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 $u \in C^\infty(\bar{\Omega})$ 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 [Str14, p. 40]).** *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.*

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 $\Omega$. 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 $\Omega$. 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 $\varphi$ 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

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

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, for example 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 smooth manifold.*

**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**.*

One should think of a state $(x, v)$ of a configuration space as follows: $x$ gives the position of the mechanical system and $v$ its velocity. The fundamental principle governing motions of mechanical systems is the following.

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

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 a smooth manifold $M$ and a function $L \in C^\infty(TM \times \mathbb{R})$, called a **Lagrangian function**.*

**Example 1.8.** For a smooth manifold $M$ 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 $M$ be a smooth manifold, $x_0, x_1 \in M$ and $t_0, t_1 \in \mathbb{R}$ with $t_0 \leq t_1$. Define the **path space of $M$ connecting $(x_0, t_0)$ and $(x_1, t_1)$** to be the set*

$$\mathcal{P}(M)_{x_1,t_1}^{x_0,t_0} := \{\gamma \in C^\infty([t_0, t_1], M) : \gamma(t_0) = x_0 \text{ and } \gamma(t_1) = x_1\}. \tag{5}$$

**Remark 1.10.** For the sake of simplicity, we will just use the terminology *path space* for $\mathcal{P}(M)_{x_1,t_1}^{x_0,t_0}$ and simply write $\mathcal{P}(M)$.

**Definition 1.11 (Variation).** *Let $\mathcal{P}(M)$ be a path space and $\gamma \in \mathcal{P}(M)$. A **variation of $\gamma$** 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 $t \in [t_1, t_0]$.*
- *$\Gamma(t_0, \varepsilon) = x_0$ for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.*
- *$\Gamma(t_1, \varepsilon) = x_1$ for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.*

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

**Example 1.13 (Perturbation of a Path along a Single Direction).** Let $M^n$ be a smooth manifold, $(U, \varphi)$ a chart and suppose that $\gamma$ is a path in $U$. With respect to this chart, we can write the coordinate representation of $\gamma$ as

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

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

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

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

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

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

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

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

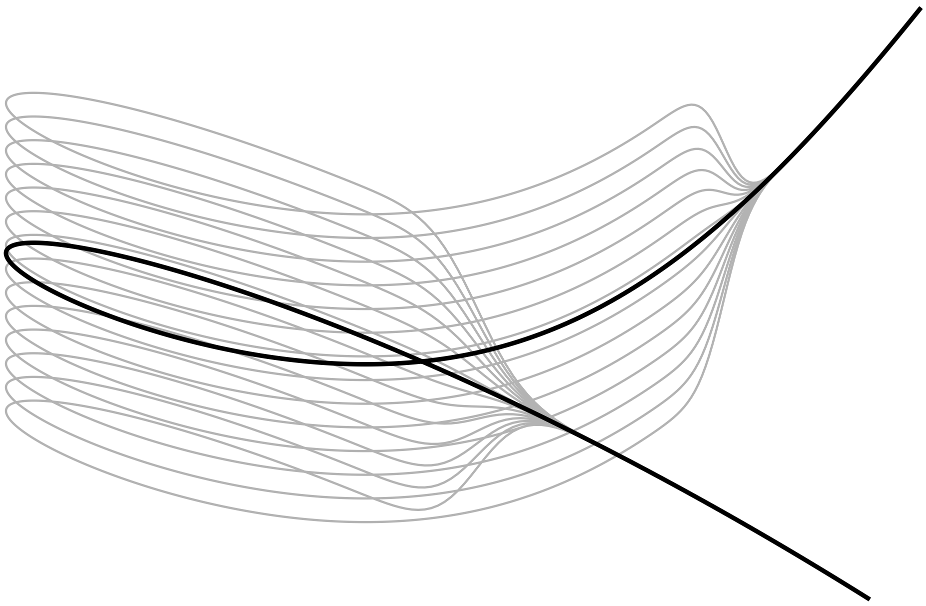


Figure 1. Example of a variation of the path $\gamma(t) = (\gamma^1(t), \gamma^2(t))$ in $\mathbb{R}^2$ defined by $\gamma(t) := (t^2 + \sin(t)\cos(t), t^3 - t)$ for $t \in [-\frac{3}{2}, \frac{3}{2}]$ along the second coordinate using a smooth bump function as in [Lee13, p. 42].

**Exercise 1.14.** Let $(X, d)$ be a metric space and $A \subseteq U \subseteq X$ where $U$ is open in $X$ and $A$ is closed in $X$. Then there exists $\delta > 0$ such that

$$U_\delta := \{x \in X : \operatorname{dist}(x, A) < \delta\} \subseteq U.$$

**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 associated to the Lagrangian system $(M, L)$***.

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 $(x_0, t_0)$ and $(x_1, t_1)$ if and only if*

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

*for all variations $\gamma_\varepsilon$ of $\gamma$.*

**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. But first, let us fix some notation. Let $M^n$ be a smooth manifold and $(U, \varphi)$ be a chart on $M$ with coordinates $(x^i)$. In what follows, we will use the abbreviation

$$\frac{\partial}{\partial x} := \left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right),$$

where as usual $\frac{\partial}{\partial x^i} : U \to TM$ denotes the *$i$-th coordinate vector field*, that is

$$\frac{\partial f}{\partial x^i}(x) := \left.\frac{\partial}{\partial x^i}\right|_x f = \partial_i (f \circ \varphi^{-1}) \left(\varphi(x)\right),$$

for all $i = 1, \dots, n$, $x \in U$ and $f \in C^\infty(M)$.

**Theorem 1.17 (Euler-Lagrange Equations).** *Let $(M^n, L)$ be a Lagrangian system. A path $\gamma \in C^\infty([t_0, t_1], M)$ describes a motion of $(M, L)$ between $(x_0, t_0)$ and $(x_1, t_1)$ if and only if with respect to all charts $(U, x^i)$*

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

*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 $\gamma$ is an extremal of the action functional $S$. The proof is divided into two steps.
*Step 1: Suppose that $\gamma$ is contained in a chart domain $U$.* Let $t \in [t_0, t_1]$ and abreviate $x_t := (\gamma(t), \dot{\gamma}(t), t)$. Suppose $\Gamma : [t_0, t_1] \times [-\varepsilon_0, \varepsilon_0] \to M$ is a variation of $\gamma$. Then there exists a rectangle $\mathcal{R}$ such that

$$[t_0, t_1] \times \{0\} \subseteq \mathcal{R} \subseteq [t_0, t_1] \times [-\varepsilon_0, \varepsilon_0]$$

and $\Gamma(\mathcal{R}) \subseteq U$. Indeed, $\Gamma$ is continuous since $\Gamma$ is smooth and so $\Gamma^{-1}(U)$ is open in $[t_0, t_1] \times [-\varepsilon_0, \varepsilon_0]$. Since $\gamma$ is a path in $U$, we get

$$[t_0, t_1] \times \{0\} \subseteq \Gamma^{-1}(U)$$

by the definition of a variation. By exercise 2.4. (c) [Lee11, p. 22], the standard Euclidean metric and the *maximum metric* $|\cdot|_\infty$ generate the same topology, thus for all $t \in [t_0, t_1]$ there exists $r_t > 0$ such that

$$B_{r_t}(t, 0) := \left\{(x, \varepsilon) \in [t_0, t_1] \times [-\varepsilon_0, \varepsilon_0] : \max\{|x - t|, |\varepsilon|\} < r_t\right\} \subseteq \Gamma^{-1}(U).$$

Since $[t_0, t_1] \times \{0\}$ is compact in $[t_0, t_1] \times [-\varepsilon_0, \varepsilon_0]$, we find $m \in \mathbb{N}$ such that

$$[t_0, t_1] \times \{0\} \subseteq \bigcup_{i=1}^{m} B_{r_i}(t_i, 0).$$

Set $r := \max_{i=1,\dots,m} r_i$ and define $\mathcal{R} := [t_0, t_1] \times (-r, r)$. Then if $(t, \varepsilon) \in \mathcal{R}$ we get that there exists some index $i$ such that $(t, 0) \in B_{r_i}(t_i, 0)$. Hence $|t - t_i| < r_i$ and so

$$|(t, \varepsilon) - (t_i, 0)|_\infty = \max\{|t - t_i|, |\varepsilon|\} < r_i.$$

Thus $(t, \varepsilon) \subseteq B_{r_i}(t_i, 0) \subseteq \Gamma^{-1}(U)$ and so $\Gamma(\mathcal{R}) \subseteq U$. Hence we can write

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

and

$$\dot{\gamma}_\varepsilon(t) = \left(\dot{\gamma}_\varepsilon^1(t), \dots, \dot{\gamma}_\varepsilon^n(t)\right)$$

for all $(x, \varepsilon) \in \mathcal{R}$, where the dot denotes a derivative with respect to time.
Using the formula for the derivative of a function along a curve A.1, we compute

$$\begin{aligned}
\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} L\left(\gamma_\varepsilon(t), \dot{\gamma}_\varepsilon(t), t\right) &= dL_{x_t}\left(\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \gamma_\varepsilon(t), \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \dot{\gamma}_\varepsilon(t), 0\right) \\
&= dL_{x_t}\left(\frac{d\gamma_\varepsilon^j(t)}{d\varepsilon}(0) \left.\frac{\partial}{\partial x^j}\right|_{\gamma(t)}, \frac{d\dot{\gamma}_\varepsilon^j(t)}{d\varepsilon}(0) \left.\frac{\partial}{\partial v^j}\right|_{\dot{\gamma}(t)}, 0\right).
\end{aligned}$$

for all variations $\gamma_\varepsilon$ of $\gamma$ in $U$. Moreover, using the formula for the differential of a function on coordinates (A.1) yields

$$dL_{x_t} = \frac{\partial L}{\partial x^i}(x_t) dx^i|_{x_t} + \frac{\partial L}{\partial v^i}(x_t) dv^i|_{x_t} + \frac{\partial L}{\partial t}(x_t) dt|_{x_t}.$$

Therefore

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

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 $\gamma_\varepsilon$ be the variation of $\gamma$ defined in example 1.13 along the $j$-th direction. Above computation therefore yields

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

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

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

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

Conversly, if we assume that the Euler-Lagrange equations (7) hold, above computation yields

$$
\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} S(\gamma_\varepsilon) = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x^i}(x_t) - \frac{d}{dt}\frac{\partial L}{\partial v^i}(x_t)\right)\frac{d\gamma_\varepsilon^i(t)}{d\varepsilon}(0)dt = 0
$$

for every variation $\gamma_\varepsilon$ of $\gamma$.

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

**Lemma 1.18 (Lebesgue Number Lemma [Lee11, p. 194]).** *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 $\delta$ is contained in a member of the family.*

Let $(U_\alpha)_{\alpha\in A}$ be the smooth structure on $M$, i.e. the maximal smooth atlas. Since $\gamma$ 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 $N \in \mathbb{N}$ such that $(t_1 - t_0)/N < \delta$ and define

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

for all $i = 0, \dots, N$. Then for all $i = 1, \dots, N$, $\gamma|_{[t_{i-1}, t_i]}$ is contained in $U_\alpha$ for some $\alpha \in A$. Let us extend the construction of example 1.13. Suppose $f \in C_c^\infty(t_{i-1}, t_i)$. Then we can define a variation $\Gamma : [t_0, t_1] \times [-\varepsilon_0, \varepsilon_0] \to M$ as follows: Define

$$\Gamma : ([t_0, t_1] \setminus \operatorname{supp} f) \times [-\varepsilon_0, \varepsilon_0] \to M$$

by $\Gamma(t, \varepsilon) := \gamma(t)$, and $\Gamma : (t_{i-1}, t_i) \times [-\varepsilon_0, \varepsilon_0] \to M$ to be the map defined in example 1.13. Since both definitions agree on the overlap $(t_{i-1}, t_i) \setminus \operatorname{supp} f$, an application of the gluing lemma for smooth maps [Lee13, p. 35] yields the existence of a variation $\Gamma$ of $\gamma$ on $M$. Therefore, step 1 implies the Euler-Lagrange equations (7). The converse direction is content of problem 1-1. □

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

**Example 1.20 (Motions on Riemannian Manifolds).** Let $(M^n, g)$ be a Riemannian manifold and consider the Lagrangian $L$ on $M$ defined in example 1.8 with kinetic energy

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

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

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

where $g_{ij}(x) := g_x\left(\frac{\partial}{\partial x^i}\big|_x, \frac{\partial}{\partial x^j}\big|_x\right)$. Thus

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

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

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

implies

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

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]). The system of equations (8) is called ***geodesic equations*** (see [Lee97, p. 58]). Hence extremals $\gamma$ of the action functional satisfy the geodesic equation and are therefore geodesics on the Riemannian manifold $M$.

**Lemma 1.21.** *Let $(M, L)$ be a Lagrangian system and define $L + df \in C^\infty(TM \times \mathbb{R})$ by*

$$(L + df)(x, v, t) := L(x, v, t) + df_x(v)$$

*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 $\gamma_\varepsilon$ is a variation of $\gamma$ in $M$. Using the formula for the derivative of a function along a curve [Lee13, p. 283] we compute

$$\begin{aligned}
\widetilde{S}(\gamma_\varepsilon) &= \int_{t_0}^{t_1} L(\gamma_\varepsilon(t), \dot{\gamma}_\varepsilon(t), t)dt + \int_{t_0}^{t_1} df_{\gamma_\varepsilon(t)}\left(\dot{\gamma}_\varepsilon(t)\right) dt \\
&= S(\gamma_\varepsilon) + \int_{t_0}^{t_1} (f \circ \gamma_\varepsilon)'(t)dt \\
&= S(\gamma_\varepsilon) + f\left(\gamma_\varepsilon(t_1)\right) - f\left(\gamma_\varepsilon(t_0)\right) \\
&= S(\gamma_\varepsilon) + f(x_1) - f(x_0).
\end{aligned}$$

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 1.22.** Lemma 1.21 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 1.23 (Autonomous System).** *A Lagrangian system $(M, L)$ is said to be an* ***autonomous Lagrangian system****, iff $L \in C^\infty(TM)$.*

Let $(M^n, L)$ be an autonomous Lagrangian system and $(U, x^i)$ a chart on $M$. Moreover, let $(x^i, v^i)$ denote standard coordinates on $TM$, that is $v^i := dx^i$ for all $i = 1, \ldots, n$. Expanding the Euler-Lagrange equations (7) yields

$$\begin{aligned}
\frac{\partial L}{\partial x^j}\left(\gamma(t), \dot{\gamma}(t)\right) &= \frac{d}{dt}\frac{\partial L}{\partial v^j}\left(\gamma(t), \dot{\gamma}(t)\right) \\
&= \frac{\partial^2 L}{\partial x^i \partial v^j}\left(\gamma(t), \dot{\gamma}(t)\right)\dot{\gamma}^i(t) + \frac{\partial^2 L}{\partial v^i \partial v^j}\left(\gamma(t), \dot{\gamma}(t)\right)\ddot{\gamma}^i(t)
\end{aligned}$$

for all $j = 1, \ldots, n$. In order to solve above system of second order ordinary differential equations for $\ddot{\gamma}^i(t)$ and all initial conditions in the chart on $TU$, the matrix $\mathscr{H}_L(x, v)$

defined by

$$\mathcal{H}_L(x,v) := \left(\frac{\partial^2 L}{\partial v^i \partial v^j}(x,v)\right)^i_j \tag{9}$$

must be invertible on $TU$.

**Definition 1.24 (Nondegenrate System).** *An autonomous Lagrangian system* $(M, L)$ *is said to be* ***nondegenerate****, iff for all coordinate charts* $U$ *on* $M$*,* $\det\mathcal{H}_L(x,v) \neq 0$ *holds on* $TU$*.*

**Example 1.25 (Nondegenrate System on a Riemannian Manifold).** Let $(M, g)$ be a Riemannian manifold. Consider the Lagrangian $T-V$ with kinetic energy $T \in C^\infty(TM)$ defined by $T(v) := \frac{1}{2}|v|^2$ and potential energy $V \in C^\infty(M)$. Then the computation performed in example 1.20 yields

$$\mathcal{H}_{T-V}(x,v) = \left(g_{ij}(x)\right)^i_j$$

on every chart since $\frac{\partial V}{\partial v^i} = 0$ for every $i$, and so this Lagrangian system is nondegenerate.

The nondegeneracy of an autonomous Lagrangian system is intrinsically connected to a certain differential form in $\Omega^1(TM)$, which we will construct now. For every $(x,v) \in TM$ we can define a covector $D^{\mathcal{F}}_{(x,v)}L \in T^*_x M$ by setting

$$D^{\mathcal{F}}_{(x,v)}L := \left.\frac{\partial}{\partial v^i}\right|_{(x,v)}(L)dx^i|_x = \frac{\partial L}{\partial v^i}dx^i. \tag{10}$$

Let $(\widetilde{U}, \widetilde{x}^i)$ be another chart on $M$ such that $U \cap \widetilde{U} \neq \varnothing$. Denote the induced coordinates on $TM$ by $(\widetilde{x}^i, \widetilde{v}^i)$. Then on $U \cap \widetilde{U}$ 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

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

or equivalently

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

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, $D^{\mathcal{F}}L$ is independent of the choice of coordinates.

**Definition 1.26 (Fibrewise Differential[2]).** *Let $(M, L)$ be an autonomous Lagrangian system. The form $D^{\mathcal{F}}L \in \Omega^1(TM)$ defined on a chart $(U, x^i)$ of $M$ by*

$$D^{\mathcal{F}}L := \frac{\partial L}{\partial v^i} dx^i \tag{11}$$

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

**Remark 1.27.** The preceeding discussion showed, that the fibrewise differential $D^{\mathcal{F}}L$ is well-defined.

**Example 1.28 (Fibrewise Differential on a Riemannian Manifold).** Consider the autonomous Lagrangian system as defined in example 1.25. Then the computation performed in example 1.20 yields

$$D^{\mathcal{F}}_{(x,v)}(T - V) = g_{ij}(x)v^i dx^j$$

on every chart since $\frac{\partial V}{\partial v^j} = 0$ for all $j$.

Recall, that a 2-covector on a finite-dimensional real vector space is said to be *nondegenrate*, iff the matrix representation with respect to some basis is invertible. Moreover, a *nondegenerate 2-form* on a smooth manifold $M$ is defined to be a 2-form $\omega$, such that $\omega_x$ is a nondegenrate 2-covector for all $x \in M$ (see [Lee13, pp. 565,567]).

**Proposition 1.29.** *An autonomous Lagrangian system $(M, L)$ is nondegenerate if and only if $d(D^{\mathcal{F}}L)$ is nondegenerate.*

*Proof.* Using the computation performed in [Lee13, p. 363], we get

$$d(D^{\mathcal{F}}L) = d\left(\frac{\partial L}{\partial v^j} dx^j\right) = \frac{\partial^2 L}{\partial x^i \partial v^j} dx^i \wedge dx^j + \frac{\partial^2 L}{\partial v^i \partial v^j} dv^i \wedge dx^j.$$

Moreover, using part (e) of properties of the wedge product [Lee13, p. 356], we compute

$$d(D^{\mathcal{F}}L)\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right) = \frac{\partial^2 L}{\partial x^i \partial v^j} \det \begin{pmatrix} dx^i\left(\frac{\partial}{\partial x^k}\right) & dx^j\left(\frac{\partial}{\partial x^k}\right) \\ dx^i\left(\frac{\partial}{\partial x^l}\right) & dx^j\left(\frac{\partial}{\partial x^l}\right) \end{pmatrix}$$

[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. See also [Maz12, p. 2].

$$+\frac{\partial^2 L}{\partial v^i \partial v^j}\det\begin{pmatrix} dv^i\left(\frac{\partial}{\partial x^k}\right) & dx^j\left(\frac{\partial}{\partial x^k}\right) \\ dv^i\left(\frac{\partial}{\partial x^l}\right) & dx^j\left(\frac{\partial}{\partial x^l}\right)\end{pmatrix}$$

$$=\frac{\partial^2 L}{\partial x^i \partial v^j}(\delta^i_k\delta^j_l - \delta^i_l\delta^j_k)$$

$$=\frac{\partial^2 L}{\partial x^k \partial v^l} - \frac{\partial^2 L}{\partial x^l \partial v^k}$$

for all $k, l = 1, \dots, n$. Similarly, we compute

$$d(D^{\mathcal{F}}L)\left(\frac{\partial}{\partial v^k}, \frac{\partial}{\partial x^l}\right) = \frac{\partial^2 L}{\partial v^k \partial v^l} \qquad \text{and} \qquad d(D^{\mathcal{F}}L)\left(\frac{\partial}{\partial v^k}, \frac{\partial}{\partial v^l}\right) = 0,$$

and using skew-symmetry, we also deduce

$$d(D^{\mathcal{F}}L)\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial v^l}\right) = -\frac{\partial^2 L}{\partial v^k \partial v^l}.$$

Therefore, the matrix representing $d(D^{\mathcal{F}}L)$ with respect to the standard basis is given by the block matrix

$$d(D^{\mathcal{F}}L) = \left(\begin{array}{c|c} * & -\mathcal{H}_L \\ \hline \mathcal{H}_L & 0 \end{array}\right),$$

where $\mathcal{H}_L$ is the matrix defined in (9). Thus

$$\det\big(d(D^{\mathcal{F}}L)\big) = (-1)^n(\det\mathcal{H}_L)^2$$

Hence the matrix representation of $d(D^{\mathcal{F}}L)$ is invertible if and only if $\mathcal{H}_L$ is invertible, and the conclusion follows. □

So far, we have associated to each Lagrangian system $(M, L)$ a 1-form on $TM$, the fibrewise differential $D^{\mathcal{F}}L$. In order to get closer to our goal of dualizing the concept of a Lagrangian function, we need also a 1-form on $T^*M$. Suppose $(U, x^i)$ is a chart on $M$. The induced standard coordinates on the cotangent bundle $T^*M$ of $M$ are given by $(x^i, \xi_i)$, where $\xi_i := \frac{\partial}{\partial x^i}$, considered as an element of the double dual $T^{**}U$. On this chart, define a one 1-form $\alpha$ by $\alpha := \xi_i dx^i$. Suppose $(\widetilde{x}^i, \widetilde{\xi}_i)$ are other coordinates. Then from the computations performed at the beginning of the previous section, we have that

$$\widetilde{\xi}_i = \frac{\partial x^j}{\partial \widetilde{x}^i}\xi_j \qquad \text{and} \qquad d\widetilde{x}^i = \frac{\partial \widetilde{x}^i}{\partial x^k}dx^k.$$

Thus

$$\alpha = \widetilde{\xi}_i d\widetilde{x}^i = \frac{\partial x^j}{\partial \widetilde{x}^i}\xi_j\frac{\partial \widetilde{x}^i}{\partial x^k}dx^k = \xi_j\delta^j_k dx^k = \xi_j dx^j,$$

and so, $\alpha$ is independen of the choice of coordinates.

**Definition 1.30 (Tautological Form).** *Let $M$ be a smooth manifold. The* ***tautological form on $T^*M$****, denoted by $\alpha$, is the form $\alpha \in \Omega^1(T^*M)$ defined locally by*

$$\alpha := \xi_i dx^i,$$

*where $(x^i, \xi_i)$ denotes the standard coordinates on $T^*M$.*

**Remark 1.31.** The preceeding discussion showed, that the tautological form $\alpha$ is well-defined.

**Remark 1.32.** The tautological form $\alpha$ as well as the fibrewise derivative $D^{\mathcal{F}}L$ on an autonomous Lagrangian system $(M, L)$ admit invariant definitions, that is a coordinate free definition. For the invariant definition of $\alpha$ see [Lee13, p. 569] or [Sil08, pp. 10–11], and for the invariant definition of $D^{\mathcal{F}}L$ see [Tak08, p. 31].

**Definition 1.33 (Legendre Transform).** *A* ***Legendre transform of an autonomous Lagrangian system $(M, L)$*** *is defined to be a fibrewise mapping $\tau_L \in C^\infty(TM, T^*M)$ such that*

$$D^{\mathcal{F}}L = \tau_L^*(\alpha).$$

**Example 1.34 (Legendre Transform on a Riemannian Manifold).** Let $(M, L)$ be a Lagrangian system. Then the morphism $\tau_L : TM \to T^*M$ defined by

$$\tau_L(x, v) := \big(x, D^{\mathcal{F}}_{(x,v)}L\big) \tag{12}$$

is a Legendre transform. In particular, if we consider the Lagrangian system defined in example 1.25, we get that the above defined Legendre transform is a diffeomorphism. Indeed, suppose that $\tau_{T-V}(x, v) = \tau_{T-V}(\tilde{x}, \tilde{v})$. Then $x = \tilde{x}$ and

$$g_{ij}(x)v^i dx^j = g_{ij}(x)\tilde{v}^i dx^j$$

using example 1.28. So we must have

$$g_{ij}(x)v^i = g_{ij}(x)\tilde{v}^i$$

for all $j$. Multiplying both sides by $g^{kj}(x)$ yields $v^k = \tilde{v}^k$ for every $k$ and hence $v = \tilde{v}$. Thus $\tau_{T-V}$ is injective. Let $\xi \in T_x^*M$ be given by $\xi_i dx^i|_x$. Then $\tau_{T-V}(x, v) = (x, \xi)$, where $v$ is given in coordinates by $v^k := g^{ki}(x)\xi_i$.

Since the nondegenracy of a Lagragian system $(M, L)$ is inherently connected to the nondegenracy of the form $d(D^{\mathcal{F}}L)$ and the definition of the Legendre transform invokes the form $D^{\mathcal{F}}L$, one would expect a connection between the nondegeneracy of the Lagrangian system and a local property of Legendre transform.

**Lemma 1.35.** *A Legendre transform on a Lagrangian system is a local diffeomorphism if and only if the Lagrangian system is nondegenrate.*

*Proof.* Denote the Lagrangian system by $(M, L)$. Let $(U, x^i)$ be a chart on $M$ and denote by $(x^i, v^i)$ and $(x^i, \xi_i)$ the induced standard coordinates on $TM$ and $T^*M$, respectively. Then we compute

$$\tau_L^*(\alpha) = \tau_\alpha^*(\xi_j dx^j) = (\xi_j \circ \tau_L) d\left(x^j \circ \tau_L\right),$$

which must coincide with

$$D^{\mathcal{F}} L = \frac{\partial L}{\partial v^j} dx^j.$$

Thus in coordinates

$$\tau_L(x, v) = \left(x, \frac{\partial L}{\partial v}\right) \tag{13}$$

and so

$$d_{(x,v)}\tau_L = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & \mathcal{H}_L \end{array}\right)$$

at every $(x, v) \in TM$. Hence

$$\det\left(d_{(x,v)}\tau_L\right) = \det\mathcal{H}_L.$$

If $\tau_L$ is a local diffeomorphism, by definition, we have that some restriction of $\tau_L$ to some neighbourhood of $(x, v)$ is a diffeomorphism, and so, by properties of differentials (d) [Lee13, p. 55], we have that $d_{(x,v)}\tau_L$ is an isomorphism. Conversly, if the Lagrangian system is nondegenerate, we conclude using the inverse function theorem for manifolds [Lee13, p. 79], that $\tau_L$ is a local diffeomorphism. □

**Definition 1.36 (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 1.37 (Energy on a Riemannian Manifold).** Consider the Lagrangian system defined in example 1.25. Then the computation performed in example 1.28 yields

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

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

**Definition 1.38 (Hamiltonian Function).** *Let $(M, L)$ be an autonomous Lagrangian system and $\tau_L$ a diffeomorphic Legendre transform. The morphism $H_L \in C^\infty(T^*M)$ defined by*

$$H_L := E_L \circ \tau_L^{-1}$$

*is called the **Hamiltonian function associated to the Lagrangian function $L$**.*

**Example 1.39 (Hamiltonian function on a Riemannian Manifold).** Consider the Lagrangian system defined in example 1.25. By example 1.34 the Legendre transform $\tau_{T-V}$ is a diffeomorphism. Using example 1.37, we compute

$$\begin{aligned} H_{T-V}(x, \xi) &= E_{T-V}\left(\tau_{T-V}^{-1}(x, \xi)\right) \\ &= E_{T-V}\left(x, v\right) \\ &= T(v) + V(x) \\ &= \frac{1}{2} g_{ij}(x) v^i v^j + V(x) \\ &= \frac{1}{2} g_{ij}(x) g^{ik}(x) \xi_k g^{jl}(x) \xi_l + V(x) \\ &= \frac{1}{2} \delta_j^k \xi_j g^{jl}(x) \xi_l + V(x) \\ &= \frac{1}{2} g^{kl}(x) \xi_k \xi_l + V(x) \end{aligned}$$

where $v = \left(g^{ki}\right)_i^k \xi$.

**Theorem 1.40 (Hamilton's Equations).** *Let $\gamma$ be a motion on an autonomous Lagrangian system $(M, L)$ and suppose that $\tau_L$ is a diffeomorphic Legendre transform. Then $\gamma$ satisfies the Euler-Lagrange equations in every chart if and only if the path*

$$\left(\gamma(t), \xi(t)\right) := \tau_L\left(\gamma(t), \dot{\gamma}(t)\right)$$

*satisfies the following system of first order ordinary differential equations in every chart:*

$$\dot{\gamma}(t) = \frac{\partial H_L}{\partial \xi}\left(\gamma(t), \xi(t)\right) \qquad \text{and} \qquad \dot{\xi}(t) = -\frac{\partial H_L}{\partial x}\left(\gamma(t), \xi(t)\right) \tag{14}$$

*The equations (14) are called **Hamilton's equations**.*

*Proof.* Let $\dim M = n$. First we compute $H_L$ in standard coordinates $(x^i, \xi_i)$ on $T^*M$. By (13), the Legendre transform is given by

$$\tau_L(x, v) = \left(x, \frac{\partial L}{\partial v}(x, v)\right) \tag{15}$$

in standard coordinates on $TM$. Since $\tau_L$ is a diffeomorphism by assumption, in particular it is a local diffeomorphism (see [Lee13, p. 80]). Hence by lemma 1.35, the Lagrangian system $(M, L)$ is nondegenerate. So considering $\tau_L^{-1}(x,\xi)$, we can apply the implicit function theorem [Lee13, p. 661] to obtain $v$ implicitely from the equation

$$\xi = \frac{\partial L}{\partial v}(x, v).$$

Hence in coordinates

$$H_L(x,\xi) = \left( \frac{\partial L}{\partial v^i} v^i - L(x, v) \right) \bigg|_{\xi = \frac{\partial L}{\partial v}}.$$

Therefore

$$\frac{\partial H_L}{\partial \xi^j} = \frac{\partial}{\partial \xi_j} \left( \xi_i v^i - L(x, v) \right) \big|_{\xi = \frac{\partial L}{\partial v}} = \delta_i^j v^i = v^j.$$

Hence

$$\frac{\partial H_L}{\partial \xi^j} \left( \gamma(t), \xi(t) \right) = \dot{\gamma}^j(t),$$

for all $j = 1, \ldots, n$. Moreover, we have that

$$\frac{\partial H_L}{\partial x^j} = \frac{\partial}{\partial x^j} \left( \frac{\partial L}{\partial v^i} v^i - L(x, v) \right) \bigg|_{\xi = \frac{\partial L}{\partial v}} = -\frac{\partial L}{\partial x^j}(x, v) \bigg|_{\xi = \frac{\partial L}{\partial v}},$$

and so

$$\frac{\partial H_L}{\partial x^j} \left( \gamma(t), \xi(t) \right) = -\frac{\partial L}{\partial x^j} \left( \gamma(t), \dot{\gamma}(t) \right),$$

for all $j = 1, \ldots, n$. If the Euler-Lagrange equations (7) hold, then we get

$$\frac{\partial H_L}{\partial x^j} \left( \gamma(t), \xi(t) \right) = -\frac{d}{dt} \frac{\partial L}{\partial v^j} \left( \gamma(t), \dot{\gamma}(t) \right) = -\xi_j(t),$$

and thus the Hamilton's equations (14) hold. Conversly, if we suppose that Hamilton's equations (14) hold, we get that

$$-\frac{d}{dt} \frac{\partial L}{\partial v^j} \left( \gamma(t), \dot{\gamma}(t) \right) = -\xi_j(t) = \frac{\partial H_L}{\partial x^j} \left( \gamma(t), \xi(t) \right) = -\frac{\partial L}{\partial x^j} \left( \gamma(t), \dot{\gamma}(t) \right),$$

and so the Euler-Lagrange equations (7) are satisfied. □

**Remark 1.41.** Under some reasonable assumptions on the Lagrangian system it can be shown that the Legendre transform (12) defined in example 1.34 is always a diffeomorphism. For more details see [Maz12, p. 8].

## Problems

1-1. Adopt the theory developed in the section on the *Legendre Transform* to the non-autonomous case, that is to the case of a Lagrangian system where the Lagrangian function can depend on time.

1-2. Complete the proof of theorem 1.17 about the Euler-Lagrange equations. *Hint:* Use the generalized notion of a *fibrewise differential* established in problem 1-1.

CHAPTER 2

# Hamiltonian Mechanics

APPENDIX A

# Differential Topology

## The Differential of a Function

Recall, that if $M$ is a smooth manifold and $f \in C^\infty(M)$, then with respect to any chart $(U, x^i)$ on $M$ we have that

$$df_x = \frac{\partial f}{\partial x^i}(x) dx^i|_x \tag{A.1}$$

for all $x \in U$ (see [Lee13, p. 281]).

**Proposition A.1 (Derivative of a Function along a Curve [Lee13, p. 283]).** *Suppose $M$ is a smooth manifold, $J \subseteq \mathbb{R}$ an interval, $\gamma \in C^\infty(J, M)$ a curve on $M$ and $f \in C^\infty(M)$. Then*

$$(f \circ \gamma)'(t) = df_{\gamma(t)}\left(\gamma'(t)\right)$$

*for all $t \in J$.*

# Bibliography


[Arn89] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics 60. Springer, 1989.

[Lee11] John M. Lee. *Introduction to Topological Manifolds*. Second Edition. Springer Science+Business Media, 2011.

[Lee13] John M. Lee. *Introduction to Smooth Manifolds*. Second Edition. Graduate Texts in Mathematics. Springer, 2013.

[Lee97] John M. Lee. *Riemannian Manifolds - An Introduction to Curvature*. Springer, 1997.

[Maz12] Marco Mazzucchelli. *Critical Point Theory for Lagrangian Systems*. Progress in Mathematics 293. Birkhäuser, 2012.

[Sil08] Ana Cannas da Silva. *Lectures on Symplectic Geometry*. Corrected 2nd printing. Lecture Notes in Mathematics 1764. Springer, 2008.

[Str14] Prof. Dr. Michael Struwe. "Funktionalanalysis I und II". 2014. URL: https://people.math.ethz.ch/~struwe/Skripten/FA-I-II-11-9-2014.pdf (visited on 09/16/2018).

[Tak08] Leon A. Takhtajan. *Quantum Mechanics for Mathematicians*. Vol. 95. Graduate Studies in Mathematics. American Mathematical Society, 2008.

# Index