Università degli Studi di Pisa

Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea in Matematica

The Horocycle flow from a Hamiltonian viewpoint

20 luglio 2012

Tesi di Laurea Magistrale

Candidato
Luca Asselle
asselle@mail.dm.unipi.it

Relatore
Prof. Alberto Abbondandolo
Università di Pisa

Anno Accademico 2011/2012

Introduction

The *Hamiltonian* and *Lagrangian formalisms*, taking place respectively in the cotangent and in the tangent bundle, certainly represent the most powerful approaches in the study of classical dynamics providing different tools and advantages, which may be very useful to understand the dynamical properties of the system. If the tangent space is the natural setting for the classical calcolous of variations and for Mather's and Mañé's approaches, the cotangent space is equipped with a canonical symplectic structure which allows one to use several symplectic topological tools coming from the study of Lagrangian graphs, Hofer's theory, Floer homology... A particularly fruitful approach is the so called *Hamilton-Jacobi method* (or *Weak KAM theory*) which is concerned with the study of solutions and subsolutions of the Hamilton-Jacobi equation; in a certain sense, this approach represents the functional analytical counterpart of the above mentioned variational approach.

In the present work we want to study the geodesic and the horocycle flows for the hyperbolic plane and its quotients from the Hamiltonian viewpoint; the hyperbolic plane $\mathbb{H}$ is the open upper half-plane in $\mathbb{R}^2$ endowed with the riemannian metric

$$g(x,y) \;=\; \frac{1}{y^2}\left(dx \otimes dx + dy \otimes dy\right).$$

The group of isometries of $\mathbb{H}$ can be naturally identified with $PSL(2,\mathbb{R})$; the action of $PSL(2,\mathbb{R})$ on $\mathbb{H}$ is transitive and can be lifted to a transitive action on the unit tangent bundle $T_1\mathbb{H}$. The geodesic flow of the hyperbolic plane can be seen as the Euler-Lagrange flow of the Lagrangian

$$G(z,v) \;=\; \frac{1}{2}\,\|v\|_z^2$$

where $\|\cdot\|_z$ denotes the norm on $T\mathbb{H}$ induced by the riemannian metric. Analogously, the horocycle flow can be viewed as the Euler-Lagrange flow at energy level $\frac{1}{2}$ of the Lagrangian

$$L(z,v) \;=\; \frac{1}{2}\,\|v\|_z^2 + \eta_z(v)$$

where η is a primitive of the standard area form on $\mathbb{H}$. The energy value $\frac{1}{2}$ coincides with the *Mañé critical value* $c(L)$; hence there is a drastic change in the dynamic crossing this energy value. More precisely the orbits with energy lower than $\frac{1}{2}$ are closed while the Euler-Lagrange flow for supercritical energy value k is the reparametrization of the geodesic flow for a suitable Finsler metric, that converges to the hyperbolic metric for $k \to +\infty$. A big part of this work will be dedicated to determine the solutions of the Hamilton-Jacobi equations associated to G and L; using the method of the invariant Lagrangian graphs we will compute some Hamilton-Jacobi solutions for the geodesic flow and we will prove that the functions

$$u_a(x,y) \;=\; 2\cdot\arctan\left(\frac{x-a}{y}\right), \qquad a \in \mathbb{R}$$

are, with the constant functions, the only solutions of the Hamilton-Jacobi equation associated to L at level $\frac{1}{2}$. Furthermore, we will get the following interesting result, which represents the viceversa of the general *Hamilton-Jacobi theorem* 3.1.2

Theorem 1. *Let H be the Hamiltonian associated to L above and let ω be a 1-form on $\mathbb{H}$ such that its graph G_ω is invariant and contained in the energy level $\{H = \frac{1}{2}\}$, then ω is exact.*

This is the plan for the rest of the paper: in chapter 1 we introduce the Hamiltonian and the Lagrangian formalisms, state the connections between them and give some significant example. In chapter 2 we recall the general theory of *Mañé's critical values* for a Lagrangian L and its connections with minimizing measures, Aubry-Mather theory and coverings. In chapter 3 we examine the theory of Mañé's critical values from the Hamiltonian viewpoint; in particular we focus on the Hamilton-Jacobi equation and on the relation between its solutions and exact invariant Lagrangian graphs. In chapter 4 we recall the basic properties of the hyperbolic plane and of its geodesic and horocycle flows; we also sketch the proof of the classic theorem (Hedlund, 1936) regarding the denseness of horocycle flows orbits in the case of a hyperbolic compact surface. In chapter 5, which represents the original part of this paper, we finally study in details the horocycle and geodesic flows from the Hamiltonian viewpoint, with particular attention to the solutions of the corresponding Hamilton-Jacobi equations.

Contents

Chapter 1

Preliminaries

1.1 Lagrangian Dynamics

The aim of this chapter is to give an overwiev on the basic facts that the reader should know in order to understand what we are going to do in the next chapters. From now on M will be a boundaryless n-dimensional connected and complete riemannian manifold (we will write simply manifold); we do not suppose for the moment that M is compact. We will see step by step whether the compactness hypothesis will be necessary or not and, in the first case, we will discuss what we have to change in order to extend the theory also to the non compact case.

Definition 1.1.1. *Given a manifold M an (autonomous)* ***uniformly convex Lagrangian*** *on M is a smooth function $L : TM \longrightarrow \mathbb{R}$ satisfying the following conditions:*

1. ***Uniform convexity****: The Hessian $\frac{\partial^2 L}{\partial v_i \partial v_j}(x, v)$, calculated in linear coordinates on the fiber T_xM, is uniformly positive definite for all $(x, v) \in TM$, i.e. there is $a > 0$ such that*

$$w \cdot L_{vv}(x, v) \cdot w \geq a|w|^2 \qquad \forall\, (x, v) \in TM, \quad \forall\, w \in T_xM\,.$$

2. ***Superlinearity****: $\lim_{|v|\to+\infty} \frac{L(x,v)}{|v|} = +\infty$ uniformly on $x \in M$; equivalently, for all $a \in \mathbb{R}$ there exists $b \in \mathbb{R}$ such that*

$$L(x, v) \geq a|v| - b \qquad \forall\, (x, v) \in TM\,.$$

3. ***Boundedness****: For all $r \geq 0$ the following inequalities hold*

$$l(r) := \sup_{\substack{(x,v) \in TM \\ |v| \leq r}} L(x, v) < +\infty\,, \tag{1.1}$$

$$g(r) := \sup_{\substack{|(x,v)| \leq r \\ |w| \leq r}} w \cdot L_{vv}(x, v) \cdot w < +\infty\,, \tag{1.2}$$

Let L be a lagrangian; we will use equivalently the notations

$$\frac{\partial L}{\partial x} = L_x\,, \qquad \frac{\partial L}{\partial v} = L_v$$

for the partial derivatives of L with respect to x and v. The *Euler-Lagrange equation* associated to L is, in local coordinates, given by

$$\frac{d}{dt}\frac{\partial L}{\partial v}(x, \dot{x}) = \frac{\partial L}{\partial x}(x, \dot{x})\,. \tag{E-L}$$

This is a second order differential equation on M and it is clearly equivalent to

$$\frac{\partial^2 L}{\partial v^2}(x,\dot{x})\cdot\ddot{x} = \frac{\partial L}{\partial x}(x,\dot{x}) - \frac{\partial^2 L}{\partial v \partial x}(x,\dot{x})\dot{x}\,;$$

hence the convexity hypothesis (L_{vv} invertible) implies that (E-L) can also be seen as a first order differential equation on TM

$$\begin{cases} \dot{x} = v; \\ \\ \dot{v} = (L_{vv})^{-1}(L_x - L_{vx}\cdot v); \end{cases}$$

In other words the convexity hypothesis (also called the *Legendre condition*) allows to define a vector field X_L on TM, called the *Euler-Lagrange vector field*, such that the solutions of

$$\dot{u}(t) = X_L(u(t))\,, \qquad u(t) = (x(t),\dot{x}(t))$$

are precisely the curves satisfying the Euler-Lagrange equation. The flow φ_t of the Euler-Lagrange vector field X_L is called the *Euler-Lagrange flow*; as we will see in proposition 1.1.3 the flow φ_t is complete, i.e. every maximal integral curve of the vector field X_L has all $\mathbb{R}$ as its domain of definition. Observe that X_L is of the form

$$X_L(x,v) = (v,\cdot)$$

Suppose now that L is a given Lagrangian defined on TM and denote by $C^k([0,T],M)$ the set of all C^k-differentiable curves $\gamma : [0,T] \longrightarrow M$; then the *action* of $\gamma \in C^k([0,T],M)$ is defined by

$$\mathbb{A}_L(\gamma) := \int_0^T L(\gamma(t),\dot{\gamma}(t))\,dt\,.$$

Fixed two point $q_1, q_2 \in M$ and a positive number $T > 0$ denote by $C^k(q_1,q_2;T)$ the set of all C^k-differentiable curves $\gamma : [0,T] \longrightarrow M$ such that $\gamma(0) = q_1$ and $\gamma(T) = q_2$; the following proposition states the connection between critical points of the action functional $\mathbb{A}_L(\cdot)$ and solutions of (E-L). More precisely the differentiable critical points of $\mathbb{A}_L(\cdot)$ are solutions of (E-L) and viceversa.

Remark. *Although in the case of autonomous lagrangians "action-minimizing curves" (i.e. curves that minimize the action, we shall talk more precisely about that in chapter 2) are always differentiable, this is no longer true in the time-periodic case $L : TM \times \mathbb{T} \longrightarrow \mathbb{R}$. In this case a further condition on the lagrangian is needed*

*4. **Completeness:** The Euler-Lagrange flow is complete, i.e. every maximal integral curve of the vector field X_L has all $\mathbb{R}$ as its domain of definition.*

Indeed in the non-autonomous case the Euler-Lagrange flow in general is not complete; without the completeness hypothesis one can construct, as shown by Ball and Mizel in [5], examples of minimizers that are not C^1 and hence are not solutions of (E-L). The role of the completeness hypothesis can be explained as follows; it is possible to prove, under the above conditions, that action minimizing curves not only exist and are absolutely continuous, but they are C^1 on an open and dense full measure subset of the interval in which they are defined. Moreover they satisfy the Euler-Lagrange equation on this set, while their velocity goes to infinity on the exceptional set on which they are not C^1; asking the flow to be complete, therefore, implies that minimizers are C^1 everywhere and that they are actually solutions of (E-L). Hereafter we shall consider only the autonomous case (i.e. no dipendence on time in the lagrangian); this choice has been made only to make the discussion easier and to avoid technical issues that would be otherwise involved. However all the theory that we are going to describe in this and in the following chapters can be generalized with "small" modifications to the non-autonomous time periodic case.

Proposition 1.1.2. *If a curve $\gamma \in C^2(q_1,q_2;T)$ is a critical point of the action functional $\mathbb{A}_L$ on $C^2(q_1,q_2;T)$ then γ satisfies the Euler-Lagrange equation*

$$\frac{d}{dt}L_v(\gamma(t),\dot{\gamma}(t)) = L_x(\gamma(t),\dot{\gamma}(t))\,.$$

Proof. We can suppose that there exists a coordinate system $\{x_1, ..., x_n\}$ about $\gamma(t)$; if $h(t) \in C^k(0,0;T)$ then for every $\epsilon > 0$ sufficiently small the curve

$$\gamma_\epsilon := \gamma + \epsilon h \in C^k(q_1, q_2; T)$$

is contained in the coordinate system. The curve γ is a critical point of the action functional $\mathbb{A}_L$; this implies that the function $g(\epsilon) := \mathbb{A}_L(\gamma_\epsilon)$ has a critical point in $\epsilon = 0$ and

$$\begin{aligned}
\lim_{\epsilon \to 0} \frac{g(\epsilon) - g(0)}{\epsilon} &= \lim_{\epsilon \to 0} \int_0^T \frac{L(\gamma + \epsilon h, \dot\gamma + \epsilon \dot h) - L(\gamma, \dot\gamma)}{\epsilon} \, dt = \\
&= \int_0^T \lim_{\epsilon \to 0} \frac{\epsilon L_x \cdot h + \epsilon L_v \cdot \dot h + o(\epsilon)}{\epsilon} \, dt = \int_0^T \left[L_x \cdot h + L_v \cdot \dot h \right] dt \; = \\
&= L_v \cdot h \Big|_0^T + \int_0^T \left[L_x - \frac{d}{dt} L_v \right] \cdot h \, dt \; = \int_0^T \left[L_x - \frac{d}{dt} L_v \right] \cdot h \, dt \, .
\end{aligned}$$

Hence for any function $h \in C^k(0,0,T)$ we have

$$0 = \int_0^T \left[L_x(\gamma(t), \dot\gamma(t)) - \frac{d}{dt} L_v(\gamma(t), \dot\gamma(t)) \right] h \, dt$$

and this implies that $\gamma(t)$ satisfies (E-L). □

Remark. *If we add a closed 1-form ω to the lagrangian L, the new lagrangian $L + \omega$ also satisfies the properties $1 - 3$ and has the same Euler-Lagrange equation as L. Indeed the action functional $\mathbb{A}_{L+\omega}$ in a neighbourood U of a curve $\gamma \in C^k(q_1, q_2; T)$ satisfies*

$$\mathbb{A}_{L+\omega}(\eta) = \mathbb{A}_L(\eta) + \int_\gamma \omega \, , \qquad \forall \, \eta \in U$$

because the curve η is homologous to γ; however, since the values of $\mathbb{A}_{L+\omega}$ and $\mathbb{A}_L$ are different, minimizers of these two actions may be different. As we will see better in chapter 2 adding an exact 1-form to the lagrangian does not change the minimizers; this allows to define the Mather α and β functions, whose properties are yet quite far from be completely understood.

Given a lagrangian $L \in C^\infty(TM)$, the *energy function* associated to L is $E : TM \longrightarrow \mathbb{R}$ defined by

$$E(x,v) := \frac{\partial L}{\partial v}(x,v) \cdot v - L(x,v) \qquad \forall (x,v) \in TM. \tag{1.3}$$

Observe that E is an integral, i.e. an invariant function, for the lagrangian flow φ_t; indeed if $x(t)$ satisfies the Euler-Lagrange equation (E-L) then

$$\frac{d}{dt} E(x, \dot x) = \frac{d}{dt} \left(\frac{\partial L}{\partial v} \right) \cdot \dot x + \frac{\partial L}{\partial v} \cdot \dot v - \frac{\partial L}{\partial x} \cdot \dot x - \frac{\partial L}{\partial v} \cdot \dot v = \left[\frac{d}{dt} \left(\frac{\partial L}{\partial v} \right) - \frac{\partial L}{\partial x} \right] \cdot \dot x = 0$$

With an analogous computation one can prove that the energy is invariant if and only if the lagrangian does not dipend on the time, indeed in the non autonomous case

$$\frac{d}{dt} E(x, \dot x, t) \; = \; \left[\frac{d}{dt} \left(\frac{\partial L}{\partial v} \right) - \frac{\partial L}{\partial x} \right] \cdot \dot x - \frac{\partial L}{\partial t}(x, \dot x, t) \; = -\frac{\partial L}{\partial t}$$

which is non zero. Since we consider only the autonomous case, the energy is constant along the motions; its level sets are called *energy levels* and are invariant under φ_t. Moreover the convexity implies that the function $f(s) := E(x, sv)$ has a minimum in $s = 0$ for all $v \in T_x M$ because

$$\begin{aligned}
f'(s) &= \frac{d}{ds} E(x, sv) \; = \; \frac{d}{ds} \left[\frac{\partial L}{\partial v}(x, sv) \cdot v - L(x,v) \right] \; = \\
&= s \cdot \left(v \cdot \frac{\partial^2 L}{\partial v^2}(x, sv) \cdot v \right) + \frac{\partial L}{\partial v}(x, sv) \cdot \dot v - \frac{\partial L}{\partial v}(x, sv) \cdot \dot v \; = \; s \left(v \cdot \frac{\partial^2 L}{\partial v^2}(x, sv) \cdot v \right)
\end{aligned}$$

and hence $f'(s) < 0$ if $s < 0$ and $f'(s) > 0$ if $s > 0$. Thus

$$\min_{v\in T_xM} E(x,v) \;=\; E(x,0) \;= -L(x,0)\,;$$

by the superlinearity

$$e_0 := \max_{x\in M} E(x,0) \;= -\min_{x\in M} L(x,0) \;> -\infty \tag{1.4}$$

and then $e_0 = \min\,\{k \in \mathbb{R} \mid \pi : E^{-1}(k) \longrightarrow M \text{ is surjective}\}$[1]. Moreover since by the uniform convexity we get that

$$A := \inf_{(x,v)\in TM,\ |w|=1} w\cdot L_{vv}(x,v)\cdot w > 0$$

using hypothesis (1.1) we obtain the following estimate for the energy function

$$\begin{aligned}
E(x,v) &= E(x,0) + \int_0^{|v|} \frac{d}{ds}E\left(x, s\frac{v}{|v|}\right) ds \geq \\
&\geq \min_{x\in M} E(x,0) + \int_0^{|v|} s\left[\frac{v}{|v|}\cdot L_{vv}\left(x, s\frac{v}{|v|}\right)\cdot\frac{v}{|v|}\right] ds \geq \\
&\geq -\max_{x\in M} L(x,0) + \int_0^{|v|} sA\,ds = -l(0) + \frac{1}{2}A|v|^2\,.
\end{aligned} \tag{1.5}$$

Similarly, using hypotesis (1.2), we get

$$E(x,v) \;=\; E(x,0) + \int_0^{|v|}\frac{d}{ds}E\left(x, s\frac{v}{|v|}\right) ds \;\leq\; e_0 + \int_0^{|v|} s\,g(|v|)\,ds \;=\; e_0 + \frac{1}{2}g(|v|)|v|^2 \tag{1.6}$$

Remark. *If $k\in\mathbb{R}$ and $K\subseteq M$ is compact then $E^{-1}(k)\cap T_KM$ is compact; in particular if M is compact then each energy level is compact.*

Proposition 1.1.3. *The Euler-Lagrange flow is complete.*

Proof. Suppose that $]\alpha,\beta[$ is the maximal interval of definition of the integral curve $t \longmapsto \varphi_t(v)$ and that both α and β are finite and let $k = E(v)$; since the energy $E(\varphi_t(v)) \equiv k$ is constant along the motion, by inequality (1.5) there is $a > 0$ such that

$$0 \leq |\varphi_t(v)| \leq a \qquad \forall t \in]\alpha,\beta[\,.$$

Since $\varphi_t(v)$ is of the form $(\gamma(t),\dot\gamma(t))$ we get that $\varphi_t(v)$ remains in the interior part of the compact

$$Q := \{(y,w)\in TM \mid d_M(y,x) \leq a[|\beta-\alpha|+1], |w| \leq a+1\}$$

where $x = \pi(v)$; the Euler-Lagrangian field is uniformly Lipschitz on Q and then by the theory of ordinary differential equations we can extend the interval of definition $]\alpha,\beta[$. □

1.2 Hamiltonian systems

In the study of classical dynamics it turns often very useful to watch things starting from another point of view rather than the lagrangian one, the so called *hamiltonian formalism*; as we will briefly see, this different approach brings significant benefits and semplifications in the study of some interesting classes of problems. We first introduce hamiltonian formalism in a completely independent way from lagrangian dynamics; but we will show immediately the connections between these two different approaches. As usual let M be a manifold and let T^*M be the cotangent bundle

[1] If $k < e_0$ then there exists $\tilde{x}\in M$ such that $E(\tilde{x},0) > k$ and hence $\tilde{x} \notin \pi(E^{-1}(k))$; on the other hand if $k > e_0$ then $E(x,0) < k$ for all $x\in M$ and hence there exists $v\in T_xM$ such that $E(x,v) = k$, i.e. $\pi(x,v) = x$.

of M; the cotangent bundle T^*M is naturally equipped with a structure of symplectic manifold([2]) given by the *canonical symplectic form* $\omega := d\Theta$, where Θ is the *Liouville form* on T^*M defined by

$$\Theta_p(\xi) = p(d\pi\xi) \qquad \forall \xi \in T_p(T^*M)\,.$$

Here $\pi : T^*M \longrightarrow M$ denotes the standard projection from T^*M into M. Observe that a local chart $\mathbf{x} = (x_1, ..., x_n)$ of M induces a local chart $(\mathbf{x},\mathbf{p}) = (x_1, .., x_n, p_1, .., p_n)$ of T^*M writing $\mathbf{p} \in T^*M$ as $\mathbf{p} = \sum_i p_i dx_i$. In these coordinates the forms Θ and ω are written as

$$\begin{aligned} \Theta &= \mathbf{p} \cdot d\mathbf{x} = \sum_{i=1}^{n} p_i\, dx_i; \\ \omega &= d\mathbf{p} \wedge d\mathbf{x} = \sum_{i=1}^{n} dp_i \wedge dx_i; \end{aligned}$$

In this context we can define a *hamiltonian* as a smooth function $H : T^*M \longrightarrow \mathbb{R}$ and the associated *hamiltonian vector field* X_H by([3])

$$\imath_\omega X_H = \omega(X_H, \cdot) = -dH\,. \tag{1.7}$$

In local charts the hamiltonian vector field X_H defines the differential equation

$$\begin{cases} \dot{x} = \ \ H_p; \\ \\ \dot{p} = -H_x; \end{cases}$$

where H_p and H_q are the partial derivatives of H with respect to x and p. If we denote by ψ_t the *hamiltonian flow* then ψ_t preserves H and the symplectic form ω, indeed([4])

$$\begin{aligned} \frac{d}{dt} H &= H_x \dot{x} + H_p \dot{p} = 0\,; \\ \frac{d}{dt}(\psi_t^* \omega) &= \mathcal{L}_{X_H} \omega = d\imath_{X_H}\omega + \imath_{X_H} d\omega = d(dH) + \imath_{X_H}(0) = 0\,; \end{aligned}$$

In order to show the connections between lagrangian dynamics and the hamiltonian one it is crucial the concept of *Fenchel transform* (for further details see appendix A), which is a standard tool in the study of convex functions and allows one to transform functions on a vector space into functions on the dual space. Thus from now on we will focus only on hamiltonians obtained by the Fenchel transform of a lagrangian L, i.e.

$$H(x,p) = \max_{v \in T_x M} \Big[< p, v >_x - \, L(x,v) \Big]$$

where $< \cdot, \cdot >_x$ denotes the canonical pairing between the tangent and the cotangent space. It is clear from the definition that for such a kind of hamiltonian we have the so called *Fenchel inequality*

$$< p, v >_x \leq \ L(x,v) + H(x,p) \qquad \forall\,(x,v) \in TM, \quad \forall\,(x,p) \in T^*M$$

which plays a crucial role in the study of lagrangian and hamiltonian dynamics; in particular, equality holds if and only if $p = L_v(x,v)$. Therefore one can define the *Legendre transform*

$$\mathcal{L} : TM \longrightarrow T^*M, \quad (x,v) \longmapsto \big(x, L_v(x,v)\big)$$

which is a diffeomorphism between the tangent bundle and the cotangent one (see appendix) and represents a conjugation between the Euler-Lagrange flow on TM and the Hamiltonian one on T^*M as the proposition below states. Observe that

[2] A *symplectic manifold* is a pair (M, ω) where M is a manifold and ω is a closed and non degenerating 2-form.

[3] Given a n-form ω and a vector field X on M, $\imath_\omega X$ is the $(n-1)$-form, called the *contraction of* ω *along the vector field* X, defined by $\imath_\omega X(\cdot, ..., \cdot) := \omega(X, \cdot, ..., \cdot)$.

[4] Here $\mathcal{L}_X$ is the *Lie derivative*.

$$H \circ \mathcal{L}(x,v) \;=\; < L_v(x,v), v >_x - \, L(x,v) \;=\; E$$

where E is the energy function of L defined in the previous section.

Proposition 1.2.1. *The Legendre transform $\mathcal{L} : TM \longrightarrow T^*M$ is a conjugacy between the lagrangian and the hamiltonian flows.*

Therefore one can equivalently study the Euler-Lagrange flow or the Hamiltonian-flow, obtaining in both cases information on the dynamics of the system. Each of these equivalent approaches will provide different tools and advantages, which may be very useful to understand the dynamical properties of the system. For instance, the tangent space is the natural setting for the classical calcolous of variations and for Mather's and Mañé's approaches; on the other hand, as already mentioned, the cotangent space is equipped with a canonical symplectic structure which allows one to use several symplectic topological tools, coming from the study of Lagrangian graphs, Hofer's theory, Floer homology... Moreover, a particular fruitful approach is the so called *Hamilton-Jacobi method* (or *Weak KAM theory*) which is concerned with the study of solutions and subsolutions of Hamilton-Jacobi equation. In a certain sense, this approach represents the functional analytical counterpart of the above mentioned variational approach; for further details one can refer for instance to [6] or to [7], which provide a very clear and complete introduction to the above arguments.

Remark. *Using the facts that $H = L^*$ and $L = H^*$ from proposition A.2.3 in the appendix we obtain that the boundedness condition on L is equivalent to the following:*

3.′ *$H = L^*$ is convex and superlinear.*

Proposition 1.2.2. *Two hamiltonian flows restricted to a same regular energy level are reparametrizations of each other* (5).

Proof. Suppose that $H, G : T^*M \longrightarrow \mathbb{R}$ are two hamiltonians with same energy level $H^{-1}(k) = G^{-1}(l)$ and that k, l are regular values for H and G respectively; then if $H(x,p) = k$ we have

$$\ker d_{(x,p)}H = T_{(x,p)}H^{-1}(k) = T_{(x,p)}G^{-1}(l) = \ker d_{(x,p)}G\,.$$

Thus there exists $\lambda(x,p) > 0$ such that

$$d_{(x,p)}H = \lambda(x,p) d_{(x,p)}G$$

and equation 1.7 implies that $X_H = \lambda(x,p)X_G$ when $H(x,p) = k$. □

1.3 Examples

Here we give some basic examples of lagrangians and associated hamiltonians; although the most natural setting for lagrangians is the mechanic one we start showing the particular case of *riemannian lagrangians*, that are just mechanic lagrangians with no potential. Given a riemannian metric $g =< \cdot, \cdot >$ on TM the riemannian lagrangian on TM is defined by the kynetic energy

$$L(x,v) = \frac{1}{2}\|v\|_x^2\,; \tag{1.8}$$

its Euler-Lagrangian equation (E-L) is the equation of the geodesics of g

$$\frac{D}{dt}\dot{x} \equiv 0,$$

where $\frac{D}{dt}$ denotes the covariant derivative, and its Euler-Lagrange flow is the geodesic flow. Its corresponding hamiltonian is given by

$$H(x,p) = \frac{1}{2}\|p\|_x^2\,.$$

[5] We say that an energy level $H^{-1}(k)$ is *regular* if k is a regular value of H, i.e. $dH(x,p) \neq 0$ whenever $H(x,p) = k$.

Analogously one can define *Finsler lagrangians*, that are given also by formula 1.8 but where $\|\cdot\|_x$ is a Finsler metric, i.e. a non necessarily symmetric(6) norm on T_xM which varies smoothly on M.

Mechanic lagrangians
The *mechanic lagrangians* play a key-role in the study of classical mechanics; they are given by the kynetic energy minus a *potential energy* $U : M \longrightarrow \mathbb{R}$

$$L(x,v) \;=\; \frac{1}{2}\|v\|_x^2 - U(x)$$

The Euler-Lagrangian equation associated to a mechanic lagrangian is

$$\frac{D}{dt}\dot{x} \;= -\nabla U(x)\,.$$

Here ∇U is the gradient of U with respect to the riemaniann metric g, i.e.

$$dU(x)[v] \;= <\nabla U(x), v>_x \qquad \forall\,(x,v)\in TM\,.$$

Its energy function and its hamiltonian are given by the kinetic energy plus the potential energy

$$E(x,v) = \tfrac{1}{2}\|v\|_x^2 + U(x);$$

$$H(x,p) = \tfrac{1}{2}\|p\|_x^2 + U(x);$$

Symmetric lagrangians
The *symmetric lagrangians* are a class of lagrangians which contains the riemannian and mechanic ones; they are the lagrangians which satisfy the condition

$$L(x,v) = L(x,-v) \qquad \forall\,(x,v)\in TM\,.$$

In this case the Euler-Lagrange flow is *reversible* in the sense that $\varphi_{-t}(v) = -\varphi_t(-v)$.

Magnetic lagrangians
If one add a closed 1-form ω to a lagrangian L then the Euler-Lagrange flow does not change. Indeed the solutions of the Euler-Lagrange equation are the critical points of the action functional on the space $C(x,y;T)$ of absolutely continuous curves joining x to y with fixed time interval and, since ω is closed, the new action functional differs from the previous one by adding a costant; hence they have the same critical points. But adding a non closed 1-form to a lagrangian does change the Euler-Lagrange flow; we call *magnetic lagrangian* a lagrangian of the form

$$L(x,v) \;=\; \frac{1}{2}\|v\|_x^2 + \eta_x(v) - U(x)$$

where $\|\cdot\|_x$ is a riemaniann metric, η is a non closed 1-form on M and $U : M \longrightarrow \mathbb{R}$ is a smooth function. If $Y : TM \longrightarrow TM$ is the bundle map such that

$$d\eta(u,v) \;= <Y(u), v>$$

then the Euler-Lagrange equation is given by

$$\frac{D}{dt}\dot{x} \;=\; Y_x(\dot{x}) - \nabla U(x)\,.$$

This models the motion of a particle with unity mass and unity charge under the effect of a magnetic field with Lorentz force Y and potential energy $U(x)$; the energy functional is the same as that of the mechanic lagrangians but its hamiltonian changes because of the change in the Legendre transform

$$E(x,v) = \tfrac{1}{2}\|v\|_x^2 + U(x);$$

$$H(x,p) = \tfrac{1}{2}\|p-\eta\|_x^2 + U(x).$$

6 A norm $\|\cdot\|$ is called non symmetric if $\|\lambda x\| = \lambda\|v\|$ only for $\lambda \geq 0$.

Mañé's Lagrangians

This is a particular case of Lagrangians introduced by Ricardo Mañé in [15]; if X is a C^k vector field on M with $k \geq 2$, one can embed its flow φ_t^X into the Euler-Lagrange flow at energy level 0 associated to a certain Lagrangian, namely

$$L_X(x,v) \;=\; \frac{1}{2}\,\|v - X(x)\|_x^2\;.$$

The energy function $E(x,v)$ of such Lagrangian is given by

$$\begin{aligned} E_X(x,v) &= L_v(x,v)\cdot v - L(x,v) \;=\; \big[v - X(x)\big]\cdot v - \frac{1}{2}\,\|v - X(x)\|_x^2 \;= \\ &= \frac{1}{2}\,\|v\|_x^2 - \frac{1}{2}\,\|X(x)\|_x^2\,. \end{aligned}$$

Since the energy is constant along the motion (the Lagrangian L_X is autonomous) the Euler-Lagrange orbits $(x,\dot{x})$ with 0 energy must satisfy $\|\dot{x}\|_x^2 = \|X(x)\|_x^2$. On the other hand one can easily check that the integral curves of X satisfy the Euler-Lagrange equation, indeed $L_x(x,X(x)) \equiv L_v(x,X(x)) \equiv 0$. In particular the flow $\varphi_t^{L_X}$ restricted to

$$\Gamma_X := \mathrm{Graph}\,(X) = \{(x,X(x)) \mid x \in M\}$$

(that is clearly invariant) is conjugated to the flow of X on M and the conjugation is given by the restriction to Graph (X) of the canonical projection $\pi : TM \longrightarrow M$. In other words, the diagram

$$\begin{tikzcd} \Gamma_X \arrow[r, "\varphi_t^{L_X}"] \arrow[d, "\pi"'] & \Gamma_X \arrow[d, "\pi"] \\ M \arrow[r, "\varphi_t^X"'] & M \end{tikzcd}$$

is commutative. Using the Fenchel transform we can easily compute the corresponding Hamiltonian

$$\begin{aligned} H_X(x,p) &= \max_{v\in T_xM}\Big[<p,v>_x - L_X(x,v)\Big] \;=\; \max_{v\in T_xM}\Big[<p,v>_x - \frac{1}{2}\,\|v - X(x)\|_x^2\Big] \;= \\ &= \;<p, X(x)+p>_x - \frac{1}{2}\,\|X(x)+p - X(x)\|_x^2 \;=\; \frac{1}{2}\,\|p\|_x^2 + <p, X(x)>_x\;. \end{aligned}$$

Chapter 2

Mañé's critical values

2.1 The action potential and the critical value $c(L)$

In this section we shall introduce the first critical value that we will talk about; this critical value, as we will see, implies a drastic change in the dynamic and in the behaviour of the action potential. Later in this chapter we will state its key-role in the study of action minimizing curves and its connection with Aubry-Mather theory. Thus let M be a manifold and let $L : TM \longrightarrow \mathbb{R}$ be a lagrangian on M; we shall be interested on action minimizing curves with *free* time interval. Unless otherwise stated, all the curves will be assumed to be absolutely continuous; for $x, y \in M$ let $C_T(x,y)$ be the set of all absolutely continuous curves joining x to y whose interval of definition is $[0,T]$ and let $C(x,y)$ be the union of such $C_T(x,y)$'s

$$C_T(x,y) := \big\{\gamma : [0,T] \longrightarrow M \mid \gamma(0) = x,\ \gamma(T) = y\big\}, \qquad C(x,y) \ = \bigcup_{T>0} C_T(x,y)\,.$$

For each $k \in \mathbb{R}$ and for any $T > 0$ we can define the *T-time potential* by

$$\Phi_k(x,y;T) := \inf_{\gamma \in C_T(x,y)} \mathbb{A}_{L+k}(\gamma) \tag{2.1}$$

so that the curves which realize $\Phi_k(x,y;T)$ are the Tonelli minimizers (see section 2.2) on $C_T(x,y)$; furthermore we can define the *action potential* $\Phi_k : M \times M \longrightarrow \mathbb{R} \cup \{-\infty\}$

$$\Phi_k(x,y) := \inf_{\gamma \in C(x,y)} \mathbb{A}_{L+k}(\gamma) \quad = \inf_{\gamma \in C(x,y)} \int_0^{T(\gamma)} \Big[L(\gamma(t),\dot\gamma(t)) + k\Big]\, dt\,. \tag{2.2}$$

Observe that if there exists a closed curve γ on M with negative $(L+k)$-action then $\Phi_k(x,y) = -\infty$ for all $x, y \in M$; indeed, fixed $x_0 \in \gamma([0,T(\gamma)])$, we can join x to y with the path

$$\gamma_n := \eta * \gamma^n * \mu\,, \quad \text{where} \ \ \eta \in C(x,x_0)\,, \quad \mu \in C(x_0,y)\,.$$

Hence we get the following estimate for $\Phi_k(x,y)$

$$\Phi_k(x,y) \le \mathbb{A}_{L+k}(\gamma_n) = \mathbb{A}_{L+k}(\eta * \gamma^n * \mu) \le \mathbb{A}_{L+k}(\eta) + n\mathbb{A}_{L+k}(\gamma) + \mathbb{A}_{L+k}(\mu) \longrightarrow -\infty\,;$$

in other words we obtain that $\Phi_k(x,y) = -\infty$ by going round γ many times. This remark suggests to define the *Mañé's critical value* $c = c(L)$ as

$$c(L) := \sup\ \big\{k \in \mathbb{R}\ \big|\ \exists \text{ closed curve } \gamma \text{ with } \mathbb{A}_{L+k}(\gamma) < 0\big\}\,.$$

Notice that the function $k \longmapsto \Phi_k(x,y)$ is increasing and that the superlinearity condition on L implies that the lagrangian is bounded below; hence there is $k \in \mathbb{R}$ such that $L + k \ge 0$ and this implies that $c(L) < +\infty$. Moreover, since the function $k \longmapsto \mathbb{A}_{L+k}(\gamma)$ is increasing for any γ,

$$c(L) = \inf\ \big\{k \in \mathbb{R}\ \big|\ \mathbb{A}_{L+k}(\gamma) \ge 0 \ \ \forall\gamma \text{ closed curve}\big\}\,.$$

Example 2.1. *Let $L : TM \longrightarrow \mathbb{R}$ be a mechanic Lagrangian, that is of the form*

$$L(x,v) \;=\; \frac{1}{2}\,\|v\|_x^2 - U(x)$$

where $U : M \longrightarrow \mathbb{R}$ is a smooth function. By the superlinearity the Lagrangian L is bounded below, that is there exists a constant $b \in \mathbb{R}$ such that $L(x,v) \geq b$ for all $(x,v) \in TM$. In particular we have $-U(x) \geq b$, that implies $U(x) \leq -b$; thus

$$\sup_{x\in M} U(x) \leq -b < +\infty\,.$$

We want to prove that $c(L) = \sup_{x\in M} U(x) =: m_U$. Indeed if $k \geq m_U$ then

$$\mathbb{A}_{L+k}[\gamma] \;=\; \int_0^{T(\gamma)} \left[\frac{1}{2}\,\|\dot\gamma(t)\|_{\gamma(t)}^2 - U(\gamma(t)) + k\right] dt \;\geq\; \int_0^{T(\gamma)} \big[k - U(\gamma(t))\big]\,dt \;\geq\; 0$$

for any closed absolutely continuous curve in M; thus $c(L) \leq m_U$. On the other hand if $k < m_U$ then there exists $\bar{x} \in M$ such that $k < U(\bar{x}) \leq m_U$ and hence if $\gamma(t) \equiv \bar{x}$ then

$$\mathbb{A}_{L+k}[\gamma] \;=\; \int_0^{T(\gamma)} \big[k - U(\bar{x})\big]\,dt \;=\; \big[k - U(\bar{x})\big]\,T \;<\; 0\,.$$

This implies that $c(L) \geq m_U$ and hence $c(L) = m_U$.

Example 2.2. *Consider the Mañé's Lagrangian L_X associated to the vector field X*

$$L_X(x,v) \;=\; \frac{1}{2}\,\|v - X(x)\|_x^2\,.$$

Since $L_X(x,v) \geq 0$ for any $(x,v) \in TM$ we get that $c(L_X) \leq 0$, indeed for any $k \geq 0$ and for any closed absolutely continuous curve in M we have

$$\mathbb{A}_{L+k}[\gamma] \;=\; \int_0^{T(\gamma)} \left[\frac{1}{2}\,\|\dot\gamma(t) - X(\gamma(t))\|_{\gamma(t)}^2 + k\right] dt \;\geq\; 0\,.$$

Observe that if $k < 0$ and γ is an integral line for X, that is $\gamma(t)$ satisfies $\dot\gamma(t) = X(\gamma(t))$, then $\mathbb{A}_{L+k}[\gamma] < 0$. Therefore if the flow φ^X has a periodic orbit or a fixed point then $c(L_X) = 0$; more generally if M is compact then $c(L_X) = 0$. Indeed if γ is an integral line for X then it admits an accumulation point; therefore we can find two arbitrarily closed points x, y in the image of γ and a curve η connecting x to y with sufficiently small $(L+k)$-action to get $\mathbb{A}_{L+k}[\gamma * \eta] < 0$.

Proposition 2.1.1. *The following hold:*

1. *If $k < c(L)$ then $\Phi_k(x,y) = -\infty$ for all $x, y \in M$; conversely, if $k \geq c(L)$ then $\Phi_k(x,y) \in \mathbb{R}$ for all $x, y \in M$.*

2. *The action potential Φ_k satisfies the triangle inequality, i.e.*

$$\Phi_k(x,z) \;\leq\; \Phi_k(x,y) + \Phi_k(y,z) \qquad \forall\, x,y,z \in M\,.$$

3. *If $k \geq c(L)$ then for all $x \in M$, $\Phi_k(x,x) = 0$.*

4. *For $k \geq c(L)$ the action potential Φ_k satisfies $\Phi_k(x,y) + \Phi_k(y,x) \geq 0$; moreover, if $k > c(L)$, then the strict inequality holds if and only if $x \neq y$.*

5. *For $k \geq c(L)$ the action potential Φ_k is Lipschitz.*

Proof. We first prove 2 for all $k \in \mathbb{R}$; observe that, since $\Phi_k(x,y) \in \mathbb{R} \cup \{-\infty\}$, the inequality in 2 makes sense for all $k \in \mathbb{R}$. Let $\gamma \in C(x,y)$ and $\eta \in C(y,z)$ be two curves joining respectively x to y and y to z; then $\gamma * \eta \in C(x,z)$ and

$$\Phi_k(x,z) \leq \mathbb{A}_{L+k}(\gamma * \eta) \leq \mathbb{A}_{L+k}(\gamma) + \mathbb{A}_{L+k}(\eta)\,.$$

Hence we obtain 2 by taking the infima on $\gamma \in C(x,y)$ and on $\eta \in C(y,z)$. Now let us prove 1; if $k < c(L)$ then there exists a closed curve γ such that $\mathbb{A}_{L+k}(\gamma) < 0$. If $z := \gamma(0)$ then

$$\Phi_k(z,z) \leq \lim_{n \to +\infty} \mathbb{A}_{L+k}(\overbrace{\gamma * \ldots * \gamma}^{n \text{ times}}) = \lim_{n \to +\infty} n\mathbb{A}_{L+k}(\gamma) = -\infty\,;$$

hence item 2 implies that for all $x, y \in M$

$$\Phi_k(x,y) \leq \Phi_k(x,z) + \Phi_k(z,z) + \Phi_k(z,y) = -\infty\,.$$

Conversely, suppose that $\Phi_k(x,y) = -\infty$ for some $k \in \mathbb{R}$ and for some $x, y \in M$; then

$$\Phi_k(x,x) \leq \Phi_k(x,y) + \Phi_k(y,x) = -\infty$$

and hence there is $\gamma \in C(x,x)$ with negative action. Then $k \leq c(L)$; now, since the set

$$\big\{k \in \mathbb{R} \;\big|\; \mathbb{A}_{L+k}(\gamma) < 0 \text{ for some closed curve } \gamma\big\}$$

is open, the hypothesis $\Phi_k(x,y) = -\infty$ actually implies that $k < c(L)$. In order to prove statement 3 observe that, fixed $k \in \mathbb{R}$, by the boundedness condition there exists $q > 0$ such that

$$|L(x,v) + k| \leq q \qquad \forall\, |v| \leq 2\,;$$

now let $\gamma : [0,\epsilon] \longrightarrow M$ be a differentiable curve with $|\dot\gamma| \equiv 1$ and $\gamma(0) = x$, then

$$\begin{aligned} \Phi_k(x,x) &\leq \Phi_k(x,\gamma(\epsilon)) + \Phi_k(\gamma(\epsilon),x) \leq \mathbb{A}_{L+k}\left(\gamma\big|_{[0,\epsilon]}\right) + \mathbb{A}_{L+k}\left(\gamma(t-\epsilon)\big|_{[0,\epsilon]}\right) = \\ &= \int_0^\epsilon \Big[L(\gamma(t),\dot\gamma(t)) + k\Big]\,dt + \int_0^\epsilon \Big[L(\gamma(t-\epsilon),\dot\gamma(t-\epsilon)) + k\Big]\,dt \leq 2q\epsilon \end{aligned}$$

and letting $\epsilon \to 0$ we get $\Phi_k(x,x) \leq 0$. On the other hand the definition of $c(L)$, together with the monotonicity of the function $k \longmapsto \Phi_k(x,x)$, implies that $\Phi_k(x,x) \geq 0$ for all $x \in M$ and hence we get the thesis. Observe now that the first part of statement 4 follows obviously from items 2 and 3; in order to prove the second part of 4, suppose that are given $k > c(L)$ and $x \neq y$ such that

$$d_k(x,y) = \Phi_k(x,y) + \Phi_k(y,x) = 0\,.$$

Let $\big\{\gamma_n : [0,T_n] \longrightarrow M\big\}_{n\in\mathbb{N}}$ be a sequence of differentiable curves joining x to y such that

$$\Phi_k(x,y) = \lim_{n\to+\infty} \mathbb{A}_{L+k}(\gamma_n)\,;$$

we claim that T_n is bounded below. Indeed, suppose that $\lim_{n\to+\infty} T_n = 0$ and fix $a > 0$; then by the superlinearity there is $b > 0$ such that $L(x,v) \geq a|v| - b$ on the whole tangent bundle and hence

$$\Phi_k(x,y) = \lim_{n\to+\infty} \int_0^{T_n} \Big[L(\gamma_n(t),\dot\gamma_n(t)) + k\Big]\,dt \geq \lim_{n\to+\infty} \left[a\int_0^{T_n} |\dot\gamma_n(t)|\,dt + (k-b)T_n\right] \geq a\, d_M(x,y)$$

and letting $a \to +\infty$ we get $\Phi_k(x,y) = +\infty$ which is false. Now let $\big\{\eta_n : [0,S_n] \longrightarrow M\big\}_{n\in\mathbb{N}}$ be a sequence of differentiable curves joining y to x and such that

$$\Phi_k(y,x) = \lim_{n\to+\infty} \mathbb{A}_{L+k}(\eta_n)\,;$$

choose $0 < T < \liminf\limits_{n\to+\infty} T_n$ and $0 < S < \liminf\limits_{n\to+\infty} S_n$. Then for $c = c(L) < k$ we have

$$
\begin{aligned}
\Phi_c(x,x) &\leq \lim_{n\to+\infty} \mathbb{A}_{L+c}(\gamma_n * \eta_n) \leq \\
&\leq \lim_{n\to+\infty} \left[\mathbb{A}_{L+k}(\gamma_n)\mathbb{A}_{L+k}(\eta_n)\right] + (c-k)(T+S) = \\
&= \Phi_k(x,y) + \Phi_k(y,x) + (c-k)(T+S) = \\
&= d_k(x,y) + (c-k)(T+S) = (c-k)(T+S) < 0
\end{aligned}
$$

which contradicts item 3. Now let us prove 5; given $k \geq c(L)$ and x_1, $x_2 \in M$ we have that

$$\Phi_k(x_1,x_2) \leq \mathbb{A}_{L+k}(\gamma) \leq q\, d_M(x_1,x_2)$$

where $\gamma : [0, d_M(x_1,x_2)] \longrightarrow M$ is a unit speed minimizing geodesic joining x_1 to x_2 and $q > 0$ is as above. If $y_1, y_2 \in M$ then the triangle inequality implies that

$$
\begin{cases}
\Phi_k(x_1,y_1) \leq \Phi_k(x_1,x_2) + \Phi_k(x_2,y_1); \\
\Phi_k(x_2,y_1) \leq \Phi_k(x_2,y_2) + \Phi_k(y_2,y_1);
\end{cases}
$$

and hence combining these two inequalities we obtain

$$\Phi_k(x_1,y_1) - \Phi_k(x_2,y_2) \leq \Phi_k(x_1,x_2) + \Phi_k(y_2,y_1) \leq q\big[d_M(x_1,x_2) + d_M(y_1,y_2)\big]$$

and finally item 5 follows changing the roles of (x_1,y_1) and (x_2,y_2). $\square$

Although the action potential Φ_k is not necessarily symmetric, the proposition above implies that the function $d_k : M \times M \longrightarrow \mathbb{R}$ defined by

$$d_k(x,y) := \Phi_k(x,y) + \Phi_k(y,x)$$

is a metric for $k > c(L)$ and a pseudo-metric for $k = c(L)$([1]). If $k \geq c(L)$ it makes sense the definition of *global minimizers* (or *time free minimizers*) for $L+k$ as a curve $\gamma \in C(x,y)$ such that

$$\mathbb{A}_{L+k}(\gamma) = \Phi_k(x,y)\,. \tag{2.3}$$

Lemma 2.1.2. *Let $\gamma : [0,T] \longrightarrow M$ be an absolutely continuous curve and let $k \in \mathbb{R}$; for $\lambda > 0$ let $\gamma_\lambda(t) := \gamma(\lambda t)$ and let $\mathcal{A}(\lambda) := \mathbb{A}_{L+k}(\gamma_\lambda)$. Then*

$$\mathcal{A}'(1) = \int_0^T \Big[E(\gamma(t),\dot\gamma(t)) - k\Big]\,dt\,. \tag{2.4}$$

Proof. Since $\dot\gamma_\lambda(t) = \lambda\dot\gamma(\lambda t)$ we get

$$\mathcal{A}(\lambda) = \mathbb{A}_{L+k}(\gamma_\lambda) = \int_0^{\frac{T}{\lambda}} \Big[L(\gamma(\lambda t), \lambda\dot\gamma(\lambda t)) + k\Big]\,dt\,;$$

hence differentiating $\mathcal{A}(\lambda)$ and evaluating at $\lambda = 1$ we have

$$\mathcal{A}'(1) = \int_0^T \Big[L_x t\dot x + L_v(\dot x + t\ddot x)\Big]\,dt - T\big[L(\gamma(T),\dot\gamma(T) + k\big]$$

and integrating by parts the term $\big(L_x\dot x + L_v\ddot x\big)t = \big(\frac{d}{dt}L\big)\,t$ we get

$$
\begin{aligned}
\mathcal{A}'(1) &= -T\big[L(\gamma(T),\dot\gamma(T)) + k\big] + Lt\Big|_0^T + \int_0^T \Big(L_v\dot x - L\Big)\,dt = \\
&= -Tk + \int_0^T E(\gamma(t),\dot\gamma(t))\,dt = \int_0^T \Big[E(\gamma(t),\dot\gamma(t)) - k\Big]\,dt\,.
\end{aligned}
$$

$\square$

[1] It could be that $d_{c(L)}(x,y) = 0$ for some $x \neq y \in M$.

Corollary 2.1.3. *A free time minimizer γ for $L + k$ has energy $E \equiv k$.*

Proof. Since γ minimizes with free time the derivative in equation (2.4) must be zero and since the energy is constant (because γ satisfies the Euler-Lagrangian equation and the Energy is a prime integral for the Euler-Lagrange flow φ_t) we get $E(\gamma, \dot{\gamma}) \equiv k$. □

We enounce now an important proposition about the *T-time potential* $\Phi_k(x, y; T)$ defined in (2.1); we refer to [6] for the proof and for further details.

Proposition 2.1.4. *The following hold:*

1. *Given a compact subset $K \subseteq M$ and $\epsilon > 0$ the function*

$$K \times K \times [\epsilon, +\infty) \longrightarrow \mathbb{R} \cup \{-\infty\}, \qquad (x, y, t) \longmapsto \Phi_k(x, y; t)$$

   is Lipschitz.

2. $\Phi_k(x, y; T) = \Phi_{c(L)}(x, y; T) + (k - c(L))T$ *for all* $k \geq c(L)$ *and for all* $x, y \in M$.

3. $\lim_{\epsilon \to 0^+} \Phi_k(x, y; \epsilon) = +\infty$ *for all* $k \geq c(L)$ *and for all* $x \neq y$.

4. $\lim_{T \to +\infty} \Phi_k(x, y; T) = +\infty$ *for all* $k > c(L)$ *and for all* $x, y \in M$.

5. $\lim_{T \to +\infty} \Phi_k(x, y; T) = -\infty$ *for all* $k < c(L)$ *and for all* $x, y \in M$.

6. *If M is compact the limits in 4 and 5 are uniform in (x, y).*

We proved above that if a global minimizer for $L + k$ exists then it has constant energy $E \equiv k$ but we do not know yet if such a minimizer effectively exists; the answer to this question turns to be affirmative in the case $k > c(L)$ as the proposition below shows. Observe that, since for $k < c(L)$ the action potential $\Phi_k \equiv -\infty$, there are no minimizers in the case $k < c(L)$; the case $k = c(L)$ is more complicated and it will be discussed later.

Proposition 2.1.5. *If $k > c(L)$ and $x \neq y \in M$ then there exists a global minimizer for $L + k$, i.e. an absolutely continuous curve $\gamma \in C(x, y)$ such that*

$$\mathbb{A}_{L+k}(\gamma) \;=\; \Phi_k(x, y)\,;$$

moreover, the energy of such γ is constant and equal to k.

Proof. If $f(t) := \Phi_k(x, y; t)$ then, by proposition 2.1.4, $f(t)$ is continuous and

$$\lim_{t \to 0^+} f(t) \;=\; \lim_{t \to +\infty} f(t) \;=\; +\infty\,;$$

hence it attains its minimum at some $T > 0$. Moreover

$$\Phi_k(x, y) \;=\; \inf_{t > 0} \Phi_k(x, y; t) \;=\; \Phi_k(x, y; T)$$

and hence it suffices to take a Tonelli minimizer $\gamma \in C_T(x, y)$ (see section 2.2) to get the thesis. □

We end this section showing another important property of the critical value $c(L)$, i.e. that the function $L \longmapsto c(L)$ is continous if we endow the set of lagrangians with the topology induced by the supremum norm on compact subsets of TM. This fact can be easily deduced from the following

Lemma 2.1.6. *Given a lagrangian L the function*

$$c : C^\infty(M, \mathbb{R}) \longrightarrow \mathbb{R} \cup \{-\infty\}, \quad \psi \longmapsto c(L + \psi)$$

is continuous in the topology induced by the supremum norm.

Proof. Suppose that the sequence $\{\psi_n\}_{n\in\mathbb{N}}$ converges to a function ψ in the supremum norm and define $c_n := c(L+\psi_n)$, $c := c(L+\psi)$; we want to prove that $c_n \longrightarrow c$. Fix $\epsilon > 0$; since $c-\epsilon < c$, by the definition of the critical value there exists a closed curve $\gamma : [0,T] \longrightarrow M$ such that $\mathbb{A}_{L+\psi+c-\epsilon}(\gamma) < 0$ and hence for n large enough we have

$$\mathbb{A}_{L+\psi_n+c-\epsilon}(\gamma) < 0\,.$$

Therefore for n sufficiently large we have $c-\epsilon < c_n$ and thus $c-\epsilon \le \liminf_{n\to+\infty} c_n$; since ϵ was arbitrary

$$c \;\le\; \liminf_{n\to+\infty} c_n\,.$$

By contradiction suppose now that $c < \limsup_{n\to+\infty} c_n$ and take $\epsilon > 0$ such that

$$c < c+\epsilon < \limsup_{n\to+\infty} c_n\,; \tag{2.5}$$

since $\psi_n \to \psi$ there exists $n_0 \in \mathbb{N}$ such that

$$-\epsilon \le \psi - \psi_n \le \epsilon \qquad \forall\, n \ge n_0\,. \tag{2.6}$$

At the same time by 2.5 there exists $m \ge n_0$ such that $c < c+\epsilon < c_m$; hence by the definition of critical value there exists a closed curve $\gamma : [0,T] \longrightarrow M$ such that

$$\mathbb{A}_{L+\psi_m+c+\epsilon}(\gamma) < 0$$

and thus, using 2.6,

$$\mathbb{A}_{L+\psi+c}(\gamma) \le \mathbb{A}_{L+\psi_m+c+\epsilon}(\gamma) < 0$$

which yields a contradiction to the definition of the critical value c. □

2.2 Tonelli's theorem

The aim of this section is to prove the important *Tonelli's theorem*, used in the proof of proposition 2.1.5, that assures the existence of Tonelli minimizers for any $k \in \mathbb{R}$ and for any $T > 0$; we do not give all the details of the proof, the reader can refer to [6] for a complete proof. Before enouncing the theorem we need to introduce some basic definitions and notations; given $T > 0$ and $x, y \in M$ let $C_T(x,y)$ be the set of all absolutely continuous curves $\gamma : [0,T] \longrightarrow M$ joining x to y and let

$$\Phi_0(x,y;T) \;=\; \inf_{\gamma\in C_T(x,y)} \mathbb{A}_L(\gamma)\,.$$

Observe that $\Phi_0(x,y;T) > -\infty$ because by the superlinearity L is bounded below; moreover if $\gamma \in C^{ac}([0,T],M)$ we can define

$$S^+(\gamma) := \mathbb{A}_L(\gamma) - \Phi_0(\gamma(0),\gamma(T);T)\,.$$

The absolutely continuous curves γ such that $S^+(\gamma) = 0$ are called *Tonelli minimizers*; observe that a Tonelli minimizer is a solution of (E-L), indeed if $x = \gamma(0)$ and $y = \gamma(T)$

$$\mathbb{A}_L(\gamma) \;=\; \Phi_0(\gamma(0),\gamma(T);T) \;=\; \inf_{\eta\in C_T(x,y)} \mathbb{A}_L(\eta)$$

and hence γ is a critical point for the action functional.

Theorem 2.2.1 (Tonelli)**.** *Let L be a lagrangian on M, then for all $x, y \in M$ and for all $T > 0$ there exists a Tonelli minimizer in $C_T(x,y)$.*

The main idea of Tonelli's theorem is to prove that the sets

$$\mathcal{A}(b) := \{\gamma \in C_T(x,y) \mid \mathbb{A}_L(\gamma) \le b\}$$

are compact in the C^0-topology; then a Tonelli minimizer will be a curve in

$$\bigcap_{b\geq\alpha} \mathcal{A}(b) \neq \emptyset$$

where $\alpha := \inf_{\eta\in C_T(x,y)} \mathbb{A}_L(\eta) \geq \inf L > -\infty$.

Definition 2.2.2. *A family $\mathcal{F} \subseteq C^0([a,b],M)$ is* ***absolutely equicontinuous*** *if for all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\sum_{i=1}^{n} |t_i - s_i| < \delta \quad \Longrightarrow \quad \sum_{i=1}^{n} d_M(x_{s_i}, x_{t_i}) < \epsilon \qquad \forall\, x \in \mathcal{F}$$

whenever $]s_1,t_1[,...,]s_n,t_n[$ are disjoint intervals in $[a,b]$.

Remark. *An absolutely equicontinuous family is equicontinuous; moreover a uniform limit of absolutely equicontinuous functions is absolutely continuous.*

Lemma 2.2.3. *For all $b \in \mathbb{R}$ and for all $T > 0$ the family*

$$\mathcal{F}(b) := \{\gamma \in C^{ac}([0,T],M) \mid \mathbb{A}_L(\gamma) \leq c\}$$

is absolutely equicontinuous.

Proof. Since by the superlinearity the lagrangian L is bounded below, by adding a constant we may assume that $L \geq 0$; for $a > 0$ let

$$K(a) := \inf \left\{ \frac{L(x,v)}{|v|} \;\middle|\; (x,v) \in TM,\ |v| \geq a \right\}.$$

The superlinearity implies that $K(a) \longrightarrow +\infty$; given $\epsilon > 0$ let $a > 0$ be such that $\frac{b}{K(a)} < \frac{\epsilon}{2}$ and

$$J := \bigcup_{i=1}^{n} [s_i, t_i], \quad E := J \cap \{|\dot{x}| < a\}$$

where $0 \leq s_1 < t_1 < ... < s_n < t_n \leq T$. Then by the definition of $K(a)$ we have that $L(x_s, \dot{x}_s) \geq K(a)|\dot{x}_s|$ for all $s \in E$ and hence

$$\begin{aligned} K(a)\sum_{i=1}^{n} d(x_{s_i}, x_{t_i}) &\leq K(a)\int_E |\dot{x}|\, d\lambda + K(a)\int_{J\setminus E} |\dot{x}| d\lambda \leq \\ &\leq \int_E L(x,\dot{x}) d\lambda + aK(a)\lambda(J) \;\leq\; b + aK(a)\lambda(J) \end{aligned}$$

where λ is the Lebesgue measure on $[0,T]$. Therefore

$$\sum_{i=1}^{n} d(x_{s_i}, x_{t_i}) \leq \frac{b}{K(a)} + a\lambda(J) \leq \frac{\epsilon}{2} + a\lambda(J)$$

which implies the absolute equicontinuity of $\mathcal{F}(c)$. □

Corollary 2.2.4. *For all $b \in \mathbb{R}$ and for all $T > 0$ there exists $r > 0$ such that for all $x, y \in M$*

$$\mathcal{A}(b) \;\subseteq\; C^{ac}\big([0,T], \overline{B(x,r)}\big).$$

Theorem 2.2.5. *If the sequence $\{\gamma_n\} \subseteq \mathcal{F}(b)$ converges in the uniform topology to γ, then $\gamma \in \mathcal{F}(b)$.*

Proof. We have noticed before that a uniform limit of absolutely equicontinuous functions is absolutely continuous; we may assume that $\gamma_n([0,T])$ is contained in a compact neighbourhood K of $\gamma([0,T])$ and also that $L \geq 0$. In order to complete the proof it will be useful the following

Lemma 2.2.6. *Given K compact, $a > 0$ and $\epsilon > 0$ there exists $\delta > 0$ such that if $x \in K$, $|x-y| \leq \delta$, $|v| \leq a$ and $\omega \in \mathbb{R}^n$ then*

$$L(y, \omega) \geq L(x, v) + L_v(x, v)(\omega - v) - \epsilon\,.$$

We do not give the proof of the lemma (see, for instance, [6]); now let $\epsilon > 0$ and $E := \{|\dot{\gamma}| \leq a\}$, then by lemma 2.2.6 for n large enough we have

$$\int_E \Big[L(\gamma, \dot{\gamma}) + L_v(\gamma, \dot{\gamma})(\dot{\gamma}_n - \dot{\gamma}) - \epsilon\Big] \;\leq \int_E L(\gamma_n, \dot{\gamma}_n) \;\leq b\,. \tag{2.7}$$

It is possible to show the following claim:

$$\lim_{n \to +\infty} \int_E L_v(\gamma, \dot{\gamma})(\dot{\gamma}_n - \dot{\gamma}) = 0\,.$$

The proof of this claim is left to the reader; using the claim and letting $n \to +\infty$ in (2.7) we get

$$\int_E L(\gamma, \dot{\gamma}) - \epsilon T \leq b\,.$$

Since $E \uparrow [0, T]$ when $a \to +\infty$ and $L \geq 0$ we get

$$\int_0^T L(\gamma, \dot{\gamma}) \;=\; \lim_{a \to +\infty} \int_E L(\gamma, \dot{\gamma}) \;\leq\; b + \epsilon T$$

and hence the thesis letting $\epsilon \to 0$. □

We get finally Tonelli's theorem by proving that the sets $\mathcal{A}(b)$ are effectively compact in the C^0-topology; an addendum due to Mather (cfr. Mather [1]) shows that, with slightly stronger hypothesis, these sets are compact also in the topology of absolutely continuous curves, i.e. the topology defined by the distance

$$d_1(\gamma_1, \gamma_2) := \sup_{t \in [0,T]} d_M(\gamma_1(t), \gamma_2(t)) + \int_0^T d_{TM}([\gamma_1(t), \dot{\gamma_1}(t)], [\gamma_2(t), \dot{\gamma_2}(t)])\, dt\,.$$

Remark. *The set $\mathcal{A}(b)$ in general is not compact in the d_1-topology as the following example shows. Let $L = \frac{1}{2}|v|$ be the riemaniann lagrangian in $\mathbb{R}^2$ endowed with the flat metric and let*

$$\eta(t) := (t, 0)\,, \qquad \gamma_n(t) := (t, \frac{1}{n}\sin(2\pi n t)), \quad t \in [0, 1]\,;$$

the action $\mathbb{A}_L(\gamma_n) = 1 + 2\pi^2$ is bounded and $\gamma_n \to \eta$ in the C^0-topology. At the same time the length of γ_n is bounded below by a poligonal curve joining maxima and minima of its second component

$$l(\gamma_n) \;\geq\; \sqrt{\frac{4}{n^2} + \frac{1}{4n^2}} > 4\,;$$

hence $l(\gamma_n) \not\to l(\eta)$ and thus $\gamma_n \not\to \eta$ in the d_1-topology. Moreover, since a reparametrization preserves length, there is no reparametrization of the γ_n's which converges to η in the d_1-topology.

Theorem 2.2.7 (Mather)**.** *If L is a lagrangian on M then for all $x, y \in M$, for all $T > 0$ and $b \in \mathbb{R}$ the set $\mathcal{A}(b)$ is compact in the C^0-topology.*

Proof. By lemma 2.2.3 the family $\mathcal{A}(b)$ is equicontinuous and by corollary 2.2.4 the curves in $\mathcal{A}(b)$ have a uniform compact range; hence by Ascoli-Arzelá's theorem and by theorem 2.2.5 we obtain that the set $\mathcal{A}(b)$ is compact. □

2.3 Global minimizers in the case $k = c(L)$

In section 2.1 we proved that for $k > c(L)$ there always exists a global minimizers for $L + k$, i.e. an absolutely continuous curve $\gamma \in C(x, y)$ which satisfies

$$\mathbb{A}_{L+k}(\gamma) \;=\; \Phi_k(x, y)$$

while in the case $k < c(L)$ there are no minimizers since the action potential $\Phi_k \equiv -\infty$. Here we are interested into studying the case $k = c(L)$; hereafter we will suppose the lagrangian L fixed and we will denote simply c instead of $c(L)$. Observe that for any absolutely continuous curve $\gamma \in C(x, y)$

$$\mathbb{A}_{L+c}(\gamma) \;\geq\; \Phi_c(x, y) \;\geq\; -\Phi_c(y, x) \tag{2.8}$$

where the last inequality follows from proposition 2.1.1([2]).

Definition 2.3.1. *An absolutely continuous curve $\gamma : I \longrightarrow M$ is said **semistatic** if*

$$\mathbb{A}_{L+c}\big(\gamma\big|_{[a,b]}\big) \;=\; \Phi_c(\gamma(a), \gamma(b))$$

*for any bounded interval $[a, b] \subseteq I$, while is said **static** if*

$$\mathbb{A}_{L+c}\big(\gamma\big|_{[a,b]}\big) \;= -\Phi_c(\gamma(b), \gamma(a))$$

for any bounded interval $[a, b] \subseteq I$.

By inequality (2.8) it follows that static curves are semistatic. Moreover an absolutely continuous curve $\gamma \in C(x, y)$ is static if and only if

1. γ is semistatic.
2. $d_c(x, y) = \Phi_c(x, y) + \Phi_c(y, x) = 0$.

From corollary 2.1.3 in section 2.1 it follows immediately that

Corollary 2.3.2. *Semistatic (and in particular static) curves have energy $E \equiv c(L)$.*

Definition 2.3.3. *If $v \in TM$ we denote by $x_v(t)$ the projection $\pi\big(\varphi_t(v)\big)$ of the Euler-Lagrange flow on M where $\pi : TM \longrightarrow M$ is the canonical projection; given a lagrangian $L : TM \longrightarrow \mathbb{R}$ define*

$$\begin{aligned}
\widetilde{\mathcal{N}} = \Sigma(L) \;&:=\; \big\{w \in TM \;\big|\; x_w : \mathbb{R} \longrightarrow M \quad \textit{is semistatic}\big\}\\
\Sigma^-(L) \;&:=\; \big\{w \in TM \;\big|\; x_w :]-\infty, 0] \longrightarrow M \quad \textit{is semistatic}\big\}\\
\Sigma^+(L) \;&:=\; \big\{w \in TM \;\big|\; x_w : [0, +\infty[\longrightarrow M \quad \textit{is semistatic}\big\}\\
\mathcal{A} = \widehat{\Sigma}(L) \;&:=\; \big\{w \in TM \;\big|\; x_w : \mathbb{R} \longrightarrow M \quad \textit{is static}\big\}
\end{aligned}$$

*$\widetilde{\mathcal{N}}$ is said the **Mañé set** while $\mathcal{A}$ is called the **Aubry set**.*

We denote by $\alpha(v)$ and by $\omega(v)$ respectively the α-limit and the ω-limit of v under the Euler-Lagrange flow; using an a priori bound on the velocity $|\dot\gamma(t)|$ for a solution $\gamma \in C_T(x, y)$ of the Euler-Lagrange equation (cfr. [6], lemma 3-2.1) it is possible to prove the following

Proposition 2.3.4. *A local static curve is static, i.e. if $x_v\big|_{[a,b]}$ is static then $v \in \mathcal{A}$.*

Using the *Mather's crossing lemma* (cfr. [6], 3-8.2) one can prove the following graph property

Theorem 2.3.5 (Mañé)**.** *For all $p \in \pi(\widehat{\Sigma}(L))$ there is a unique $\xi(p) \in T_pM$ such that $(p, \xi(p)) \in \widehat{\Sigma}(L)$; moreover the map $\xi : \pi(\widehat{\Sigma}(L)) \longrightarrow \widehat{\Sigma}(L)$ is Lipschitz and $\widehat{\Sigma}(L) =$ graph ξ.*

[2]This inequality is strict if $k > c(L)$ and $x \neq y$ but in the case $k = c(L)$ the inequality may be not strict.

The canonical projection restricted to the Aubry set gives an identification $\mathcal{P} := \pi(\widehat{\Sigma}(L)) \cong \widehat{\Sigma}(L)$, where $\mathcal{P}$ is the *Peierls set*; this name is justified (see proposition below) from the fact that the Peierls set $\mathcal{P}$ is the set in which the *Peierls barrier*

$$h(x,y) := \liminf_{T\to+\infty} \ \Phi_c(x,y;T)\,; \tag{2.9}$$

takes the value 0. The difference between the action potential and the Peierls barrier is that in the Peierls barrier the curves must be defined on large time intervals.

Proposition 2.3.6. *The following hold:*

1. *The function h is Lipschitz.*
2. *For all $x, y \in M$ we have $h(x,y) \geq \Phi_c(x,y)$; in particular $h(x,x) \geq 0$ for all $x \in M$.*
3. *The Peierls barrier satisfies the triangle inequality, i.e.*

$$h(x,z) \leq h(x,y) + h(y,z) \qquad \forall\, x,y,z \in M\,.$$

4. $h(x,y) \leq \Phi_c(x,p) + h(p,q) + \Phi_c(q,y)$ *for all $x,y,p,q \in M$.*
5. $h(x,x) = 0$ *if and only if $x \in \pi(\widehat{\Sigma}(L)) = \mathcal{P}$.*
6. *If $\widehat{\Sigma}(L) \neq \emptyset$ then* $h(x,y) \ \leq \inf_{p\in\pi(\widehat{\Sigma}(L))} \big[\Phi_c(x,p) + \Phi_c(p,y)\big]$.

Proof. Item 2 is trivial; in order to prove statement 3 observe that for all $s,t>0$ and for all $y \in M$

$$h_{t+s}(x,z) \ \leq \ h_t(x,y) + h_s(y,z)\,;$$

hence taking the $\liminf_{t\to+\infty}$ we get

$$h(x,z) \ \leq \ h(x,y) + h_s(y,z) \qquad \forall\, s>0 \tag{2.10}$$

and thus item 3 follows taking the $\liminf_{s\to+\infty}$. If in (2.10) we take the infimum on $s>0$ we get

$$h(x,z) \ \leq \ h(x,y) + \Phi_c(y,z) \ \leq \ h(x,y) + A d_M(y,z) \quad \forall\, x,y,z \in M$$

where A is the Lipschitz constant for Φ_c; therefore changing the roles of x,y,z we obtain that h is Lipschitz, that is statement 1. Item 4 follows taking the $\liminf_{t\to+\infty}$ in the following inequality

$$\inf_{s>t} \ h_s(x,y) \ \leq \ \Phi_c(x,p) + h_t(p,q) + \Phi_c(q,y)$$

whose easy proof is left to the reader. Now let us focus on item 5; we prove only that if $p \in \mathcal{P}$ then $h(p,p) = 0$, the reader can refer to [6] for the rest of the proof. Take $v \in \widehat{\Sigma}$ such that $\pi(v) = p$ and $y \in \pi(\omega(v))$ where, as usual, $\omega(v)$ denotes the ω-limit of v under the Euler-Lagrange flow; let $\gamma(t) := \varphi_t(v)$ and choose a sequence $t_n \uparrow +\infty$ such that $\gamma(t_n) \to y$, then

$$\begin{aligned} 0 \ \leq \ h(p,p) \ &\leq \ h(p,y) + h(y,p) \ \leq \\ &\leq \ \lim_{n\to+\infty} \mathbb{A}_{L+c}\big(\gamma\big|_{[0,t_n]}\big) + \Phi_c(y,p) \ \leq \\ &\leq \ \lim_{n\to+\infty} -\Phi_c(\gamma(t_n),p) + \Phi_c(y,p) \ = \ 0 \end{aligned}$$

Now item 6 follows obviously from items 4 and 5. □

Corollary 2.3.7. *Since for any compact manifold M the Aubry set $\mathcal{A}$ is non empty (corollary 2.5.4), for any compact manifold the following holds*

$$h(x,y) \ = \ \inf_{p\in\pi(\widehat{\Sigma})} \Big[\Phi_c(x,p) + \Phi_c(p,y)\Big]\,.$$

The graph property in theorem 2.3.5 allows to define an equivalence relation on the Aubry set

$$u \equiv v \quad \Longleftrightarrow \quad d_c(\pi u, \pi v) = 0$$

whose equivalence classes are called *static classes.* The continuity of the pseudo-metric d_c implies that a static class is closed while from proposition 2.3.4 it follows that any static class is invariant.

Proposition 2.3.8. *If $v \in \widetilde{\mathcal{N}}$ is semistatic then $\alpha(v), \omega(v) \subseteq \mathcal{A}$; moreover $\alpha(v)$ and $\omega(v)$ are each contained in a static class.*

Proof. We prove only that $\omega(v) \subseteq \mathcal{A}$; let $\gamma(t) := \pi\varphi_t(v)$ where φ_t is the Euler-Lagrange flow and $\pi : TM \longrightarrow M$ is the canonical projection. Suppose that

$$t_n \longrightarrow +\infty\,, \qquad \dot\gamma(t_n) \longrightarrow w \in TM$$

and let $\eta(t) := \pi\varphi_t(w)$; since γ and η are solutions of the Euler-Lagangian equation we have that

$$\gamma\big|_{[t_n-s,t_n+s]} \longrightarrow \eta\big|_{[-s,s]}$$

in the C^1-norm and hence

$$\begin{aligned}
\mathbb{A}_{L+c}\big(\eta\big|_{[-s,s]}\big) + \Phi_c(\eta(s),\eta(-s)) \;&=\; \lim_{n\to+\infty}\Big[\mathbb{A}_{L+c}\big(\gamma\big|_{[t_n-s,t_n+s]}\big) + \lim_{m\to+\infty}\mathbb{A}_{L+c}\big(\gamma\big|_{[t_n+s,t_m-s]}\big)\Big] = \\
&=\; \lim_{n,m\to+\infty}\mathbb{A}_{L+c}\big(\gamma\big|_{[t_n-s,t_m-s]}\big) = \\
&=\; \lim_{n,m\to+\infty}\Phi_c(\gamma(t_n-s),\gamma(t_m-s)) \;=\; \\
&=\; \Phi_c(\eta(-s),\eta(-s)) \;=\; 0\,.
\end{aligned}$$

Thus $w \in \mathcal{A}$; now let $u \in \omega(v)$, we may assume that $\dot\gamma(s_n) \longrightarrow u$ with $t_n < s_n < t_{n+1}$. Then

$$\begin{aligned}
d_c(\pi w, \pi u) \;&=\; \Phi_c\big(\pi w, \pi u\big) + \Phi_c\big(\pi u, \pi w\big) = \\
&=\; \lim_{n\to+\infty}\Big[\mathbb{A}_{L+c}\big(\gamma\big|_{[t_n,s_n]}\big) + \mathbb{A}_{L+c}\big(\gamma\big|_{[s_n,t_{n+1}]}\big)\Big] = \\
&=\; \lim_{n\to+\infty}\mathbb{A}_{L+c}\big(\gamma\big|_{[t_n,t_{n+1}]}\big) \;=\; \Phi_c(\pi w, \pi w) \;=\; 0
\end{aligned}$$

and hence w and u are in the same static class. $\square$

Remark. *Let $L(x,v) = \frac{1}{2}\|v\|_x^2 - U(x)$ be a mechanic Lagrangian; then*

$$\Phi_{m_U}(x,y) \;=\; \inf_{\gamma\in C(x,y)}\int_0^{T(\gamma)}\left[\frac{1}{2}\,\|\dot\gamma(t)\|^2_{\gamma(t)} - U(\gamma(t)) + m_U\right]dt \;\geq\; 0\,.$$

Moreover $\mathbb{A}_{L+m_U}[\gamma] = 0$ if and only if $\gamma(t) \equiv \bar{x}$ with $U(\bar{x}) = m_U$; therefore $(\bar{x}, 0) \in \widehat{\Sigma}(L)$ and the curve $\gamma(t) \equiv \bar{x}$ is static. In particular if M is compact then the Aubry set is non empty (since a maximal point for U effectively exists); as we will see later in this chapter, this fact is true in general for any arbitrary Lagrangian. On the other hand if $x \neq y$ and $\gamma \in C(x,y)$ then

$$\mathbb{A}_{L+m_U}[\gamma] \;\geq\; \int_0^{T(\gamma)}\frac{1}{2}\,\|\dot\gamma(t)\|^2_{\gamma(t)}\,dt \;\geq\; E[\eta] \;>\; 0$$

where η is a energy-minimizing geodesic connecting x to y. Therefore for any $x \neq y$

$$d_{m_U}(x,y) \;=\; \Phi_{m_U}(x,y) + \Phi_{m_U}(y,x) \;>\; 0$$

and hence any non constant curve can not be static. In other words, the Aubry set is composed by the points $(x,0) \in TM$ such that x is a maximal point for $U(\cdot)$ and the static orbits are the constant paths $\gamma(t) \equiv x$. Furthermore, proposition 2.3.8 implies that the α,ω-limits of any semistatic orbit (contained in the energy level $\{E = m_U\}$ by corollary 2.3.2) lie in the Aubry set; therefore the semistatic orbits are contained in the stable and unstable manifolds of the points $(x,0) \in \mathcal{A}$.

2.4 Invariance of minimizing measures

In the previous sections we spoke about minimizing orbits and their existence; here we change point of view and briefly discuss about *minimizing measures.* Although we are not going to give all the details we hope that everything will be clear and understandable to the reader, who can refer to [6] for a more precise discussion. We will see later in section 2.5 the connections between what we have done so far and what we are going to do now in this section; more precisely we will state the relation between the critical value $c(L)$ and minimizing measures. Thus let M be a manifold and let C^0_l be the set of continuous functions $f : TM \longrightarrow \mathbb{R}$ having linear growth, i.e.

$$\|f\|_l := \sup_{(x,v)\in TM} \frac{|f(x,v)|}{1+\|v\|} < +\infty\,;$$

denote by $\mathcal{M}_l$ the set of Borel probabilities μ on TM such that

$$\int_{TM} \|v\|_x \, d\mu < +\infty$$

endowed with the topology such that

$$\lim_{n\to+\infty} \mu_n = \mu \quad \Longleftrightarrow \quad \lim_{n\to+\infty} \int_{TM} f\, d\mu_n = \int_{TM} f\, d\mu \quad \forall\, f \in C^0_l\,. \tag{2.11}$$

If $(C^0_l)'$ denotes the dual of C^0_l, then $\mathcal{M}_l$ is naturally embedded in $(C^0_l)'$ and its topology coincides with that induced by the weak* topology on $(C^0_l)'$. Moreover this topology is metrizable; indeed if $\{f_n\}_{n\in\mathbb{N}} \subseteq C^0_l$ is a sequence of functions with compact support which is dense in C^0_l (in the topology of uniform convergence on compact sets of TM) then the metric $d(\cdot,\cdot)$ on $\mathcal{M}_l$ given by

$$d(\mu_1,\mu_2) := \left| \int_{TM} \|v\|\, d\mu_1 - \int_{TM} \|v\|\, d\mu_2 \right| + \sum_{n\in\mathbb{N}} \frac{1}{c_n 2^n} \left| \int_{TM} f_n\, d\mu_1 - \int_{TM} f_n\, d\mu_2 \right|$$

where $c_n := \sup_{(x,v)} |f_n(x,v)|$ induces the weak* topology on $\mathcal{M}_l$ (see [6]). Now if $\gamma : [0,T] \longrightarrow M$ is a closed and absolutely continuous curve we can define a probability measure $\mu_\gamma \in \mathcal{M}_l$ by

$$\int_{TM} f\, d\mu_\gamma := \frac{1}{T}\int_0^T f(\gamma(t),\dot\gamma(t))\, dt \qquad \forall\, f \in C^0_l\,.$$

Notice that μ_γ lies ettectively in $\mathcal{M}_l$ because of the absolute continuity which implies that

$$\int |\dot\gamma(t)|\, dt < +\infty\,.$$

We call the closure $\overline{C(M)}$ of the set $C(M)$ of such μ_γ's on $\mathcal{M}_l$ the set of *holonomic measures*; observe that this set is convex. Finally, given a lagrangian L, we define $\mathcal{M}(L)$ as the set of φ_t-invariant probabilities on TM where φ_t is, as usual, the Euler-Lagrange flow. It is easy to see in the case of an autonomous lagrangian that the set $\mathcal{M}(L)$ is non empty (cfr. [7], proposition 4.1).

Remark. *Because of the conservation of the energy along the motions, each energy level is compact and invariant under the Euler-Lagrange flow; it is a well-known result by Kryloff and Bogoliouboff [11] that a flow on a compact metric space has at least an invariant probability measure.*

Remark. *One can show the existence of invariant probability measure with finite action also in the case of non-autonomous time-periodic lagrangians; as it was originally done by Mather in [1], one can apply Kryloff and Bogoliouboff's result to a one-point compactification of TM and consider the extended lagrangian system that leaves the point at infinity fixed. The main step consists in showing that the measure provided by this construction has no atomic part supported at ∞.*

Given a probability measure on TM we define its *action* as

$$\mathbb{A}_L(\mu) := \int_{TM} L\, d\mu\,.$$

Since by the superlinearity the lagrangian is bounded below (and of course $L \notin C_l^0$), this action is well define although $\mathbb{A}_L(\mu) = +\infty$ for some $\mu \in \mathcal{M}_l$; indeed one can show that there exists a probability measure $\mu \in \mathcal{M}_l$ such that

$$\int_{TM} \|v\|_x^2 \, d\mu = +\infty \, .$$

The following proposition shows that the minimum of the action on $C(M)$ is the same as the minimum on $\overline{C(M)}$ (see [6] fot the proof).

Proposition 2.4.1. *Given $\mu \in \overline{C(M)}$ there is a sequence $\{\mu_{\eta_n}\} \subseteq C(M)$ such that $\mu_{\eta_n} \to \mu$ and*

$$\lim_{n \to +\infty} \int_{TM} L \, d\mu_{\eta_n} = \int_{TM} L \, d\mu \, .$$

The statement of the proposition is not trivial; it is easy to see that the function

$$\mathbb{A}_L : \overline{C(M)} \longrightarrow \mathbb{R}$$

is lower semi-continuous because if $L_k := \min\{L, k\}$ then $\mathbb{A}_{L_k}$ is Lipschitz with constant k, indeed

$$\left|\mathbb{A}_{L_k}(\mu) - \mathbb{A}_{L_k}(\nu)\right| \ \leq \ k \, d(\mu, \nu)$$

and since $\mathbb{A}_{L_k} \uparrow \mathbb{A}_L$ we get the lower-semicontinuity. But in general $\mathbb{A}_L$ is not continuous; it is possible to give a sequence $\{\mu_{\gamma_n}\}_{n \in \mathbb{N}} \subseteq C(M)$ such that $\mu_{\gamma_n} \longrightarrow \mu$ but

$$\liminf_{n \to +\infty} \mathbb{A}_L(\mu_{\gamma_n}) \ > \ \mathbb{A}_L(\mu)$$

for a quadratic lagrangian L. This can be made by calibrating the high speeds in γ_n so that

$$\int_{\{|v| > R\}} |v| \, d\mu_{\gamma_n} \longrightarrow 0 \, , \quad \text{but} \ \ a := \liminf_{n \to +\infty} \int_{\{|v| > R\}} L \, d\mu_{\gamma_n} \ > 0 \, .$$

The limit measure μ will have support on $\{|v| \leq R\}$ and will not see the remnant a of the action.

Theorem 2.4.2 (Mañé)**.** *The following hold:*

1. *We have the chain of inclusion $\mathcal{M}(L) \subseteq \overline{C(M)} \subseteq \mathcal{M}_l$.*
2. *If $\mu \in \overline{C(M)}$ satisfies*

   $$\mathbb{A}_L(\mu) = \min\{\mathbb{A}_L(\nu) \mid \nu \in \overline{C(M)}\}$$

   then $\mu \in \mathcal{M}(L)$.
3. *If M is compact and $a \in \mathbb{R}$ then the set $\{\mu \in \overline{C(M)} \mid \mathbb{A}_L(\mu) \leq a\}$ is compact.*

Proof. We prove only statement 3; for a complete proof the reader can see Contreras [6]. From what concerning statement 1 the inclusion $\mathcal{M}(L) \subseteq \overline{C(M)}$ follows from Birkhoff's ergodic theorem and the fact that $\overline{C(M)}$ is convex, while the second inclusion follows from the fact that $\mathcal{M}_l$ is closed as one can see taking $f = \|v\|$ in equation (2.11). Observe that item 3 implies the existence of a minimizers as in statement 2; now let us prove 3. Since $\overline{C(M)}$ is closed it is enough to prove that

$$\mathcal{A}(a) := \{\mu \in \mathcal{M}_l \mid \mathbb{A}_L(\mu) \leq a\}$$

is compact in $\mathcal{M}_l$. First we prove that $\mathcal{A}(a)$ is closed; let $k > 0$ and define $L_k := \min\{L, k\}$. Moreover let

$$\mathcal{B}_k := \left\{ \mu \in \mathcal{M}_l \ \Big| \ \int_{TM} L_k d\mu \leq a \right\} \, ;$$

since $L_k \in C^0_l$ the set $\mathcal{B}_k$ is closed in $\mathcal{M}_l$ and hence $\mathcal{A}(a)$ is closed because $\mathcal{A}(a) = \cap_{k>0} \mathcal{B}_k$ is intersection of closed sets. In order to prove the compactness consider a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}(a)$; applying the Riesz's theorem we can assume, up to take a subsequence, that there exists a measure μ on the Borel σ-algebra of TM such that

$$\int_{TM} f_i \; d\mu_n \;\longrightarrow\; \int_{TM} f_i \; d\mu \tag{2.12}$$

for every f_i in the sequence used for the definition of $d(\cdot,\cdot)$. Approximating the function 1 by the functions f_i we see that μ is a probability; moreover approximating L_k by the functions f_i we get

$$\int_{TM} L_k d\mu = \lim_{n\to+\infty} \int_{TM} L_k d\mu_n \;\le\; \liminf_{n\to+\infty} \int_{TM} L d\mu_n \;\le\; a$$

and letting $k \uparrow +\infty$, by the monotone convergence theorem, we get $\mathbb{A}_L(\mu) \le a$. Now let $b > 0$ such that $|v| \le L(x,v) + b$ for all $(x,v) \in TM$; then

$$\int_{TM} |v| d\mu \le \mathbb{A}_L(\mu) + b \le a + b < +\infty$$

and hence $\mu \in \mathcal{M}_l$. We now prove that

$$\lim_{n\to+\infty} \int_{TM} |v| \; d\mu_n \;=\; \int_{TM} |v| \; d\mu \,; \tag{2.13}$$

let $\epsilon > 0$ and choose $r > 0$ such that([3])

$$L(x,v) > \frac{a}{\epsilon}|v| \qquad \forall\, |v| > r$$

Then

$$\int_{\{|v|>r\}} |v| d\mu_n \;\le\; \frac{\epsilon}{a} \int_{\{|v|>r\}} L d\mu_n \;\le\; \frac{\epsilon}{a} \int_{TM} L d\mu_n \;\le\; \epsilon\,;$$

an anologous inequality holds also for μ because $\mathbb{A}_L(\mu) \le a$. Since from 2.12 we obtain that there is $n_0 > 0$ such that

$$\left| \int_{\{|v|\le r\}} |v| \; d\mu - \int_{\{|v|\le r\}} |v| \; d\mu_n \right| < \epsilon \qquad \forall\, n > n_0$$

combining the two last inequalities we get 2.13 and hence the thesis. □

2.5 Ergodic characterization of the critical value $c(L)$

Recall that if μ is a probability measure on TM its *action* is defined by

$$\mathbb{A}_L(\mu) := \int_{TM} L \; d\mu\,;$$

since by the superlinearity the lagrangian L is bounded below and hence the action is well defined. If $\mathcal{M}(L)$ denotes the set of all φ_t-invariant probabilities on TM then the following theorem states the connection between the critical value $c(L)$ introduced in section 2.1 and the set $\mathcal{M}(L)$.

Theorem 2.5.1 (Mañé)**.** *If M is compact then*

$$c(L) \;=\; -\min\,\{\mathbb{A}_L(\mu) \mid \mu \in \mathcal{M}(L)\}\,.$$

---
[3] By adding a costant we may assume that $L > 0$.

We will deduce the theorem from the theorem below, which also applies to the non-compact case; if M is non-compact theorem 2.5.1 may not hold (cfr. Contreras [6], 5-7). Recall that if $\gamma : [0,T] \longrightarrow M$ is a closed absolutely continuous curve then we can associate to γ a measure $\mu_\gamma \in \mathcal{M}_l$ defined by

$$\int_{TM} f \, d\mu_\gamma \;=\; \frac{1}{T}\int_0^T f(\gamma(t),\dot\gamma(t))\,dt$$

for all $f \in C^0_l$; the set $\overline{C(M)}$ is the closure of such μ_γ's in $\mathcal{M}_l$.

Theorem 2.5.2. *If L is a lagrangian on M then*

$$c(L) \;= -\inf\left\{\mathbb{A}_L(\mu) \mid \mu \in \overline{C(M)}\right\} = -\inf\left\{\mathbb{A}_L(\mu) \mid \mu \in C(M)\right\}.$$

Proof. If $\mu_\gamma \in C(M)$ then $\mathbb{A}_{L+c(L)}(\mu_\gamma) \geq 0$; hence $\mathbb{A}_L(\mu_\gamma) \geq -c(L)$ and thus

$$-c(L) \;\leq\; \inf\left\{\mathbb{A}_L(\mu) \mid \mu \in C(M)\right\} = \; \inf\left\{\mathbb{A}_L(\mu) \mid \mu \in \overline{C(M)}\right\}$$

where the last equality follows from proposition 2.4.1. Conversely if $k < c(L)$ then there exists a closed absolutely continuous curve γ on M with negative action; thus $\mu_\gamma \in C(M)$ and

$$-k \;>\; \mathbb{A}_L(\mu_\gamma) \;\geq \inf\left\{\mathbb{A}_L(\mu) \mid \mu \in \overline{C(M)}\right\}.$$

We get now the thesis letting $k \uparrow c(L)$. □

This theorem, togheter with 2.4.2, implies theorem 2.5.1; indeed by theorem 2.4.2 if a minimizing measure exists, then it is invariant under the lagrangian flow. But if M is compact we know that such a minimizer exists because of the same theorem; thus if M is compact

$$\begin{aligned} c(L) &= -\inf\left\{\mathbb{A}_L(\mu) \mid \mu \in \overline{C(M)}\right\} = \\ &= -\min\left\{\mathbb{A}_L(\mu) \mid \mu \in \overline{C(M)}\right\} = -\min\left\{\mathbb{A}_L(\mu) \mid \mu \in \mathcal{M}(L)\right\}. \end{aligned}$$

We call an holonomic measure $\mu \in C(M)$ *(globally) minimizing* if $\mathbb{A}_L(\mu) = -c(L)$. Observe that if $p : N \longrightarrow M$ is a covering with M compact and $\mathbb{L} := L \circ dp$ is the lifted lagrangian then theorems 2.5.2 and 2.4.2 imply that

$$c(\mathbb{L}) \;= -\min\left\{\mathbb{A}_L(\mu) \mid \mu \in \mathcal{M}(L) \cap \overline{dp_* C(N)}\right\} \tag{2.14}$$

by noticing that $\mathbb{A}_L(dp_*\nu) = \mathbb{A}_{\mathbb{L}}(\nu)$ for each $\nu \in \overline{C(N)}$. Here $dp_*C(N)$ is the set of probabilities μ_γ on TM, where γ is a closed absolutely continuous curve on M, whose lifts to N are closed; the compactness property on theorem 2.4.2 allows to obtain a minimum in (2.14) instead of the infimum which may be not attained in the non-compact case.

Theorem 2.5.3 (Mañé)**.** *A measure $\mu \in \mathcal{M}(L)$ is minimizing if and only if $supp\,(\mu) \subseteq \widehat{\Sigma}(L) = \mathcal{A}$.*

We do not give the proof of this theorem (see for example Contreras [6], theorem 3-6.1); we just want to point out that this theorem together with theorem 2.5.1 implies the following

Corollary 2.5.4. *If M is compact then $\mathcal{A} = \widehat{\Sigma}(L) \neq \emptyset$.*

Moreover theorem 2.5.3 together with corollary 2.3.2 in section 2.3 imply

Corollary 2.5.5 (Dias Carneiro)**.** *If μ is a minimizing measure then it is supported in the energy level $E(supp\,(\mu)) = c(L)$.*

2.6 The Aubry-Mather theory and the critical value $c_0(L)$

In this section we are going to present Aubry-Mather theory; in order to do that we shall assume that M is a compact manifold. Since by this assumption, any 1-form is in C^0_l; observe that an holonomic probability μ satisfies the conditions

$$\int_{TM} df\ d\mu = 0 \qquad \forall f \in C^\infty(M,\mathbb{R})\,. \tag{2.15}$$

Indeed let $\mu_\gamma \in C(M)$, then by definition

$$\int_{TM} df\ d\mu_\gamma = \int_0^T df(\gamma(t))[\dot\gamma(t)]\,dt = \int_0^T \frac{d}{dt}f(\gamma(t))\,dt = f(\gamma(T)) - f(\gamma(0)) = 0$$

and hence equation (2.15) follows because $\overline{C(M)}$ is the closure of the set of such μ_γ's. Then, if μ is a holonomic probability, it is well defined its *homology class* $\rho(\mu) \in H_1(M,\mathbb{R}) \cong H^1(M,\mathbb{R})^*$ by

$$< \rho(\mu), [\omega] > \ = \ \int_{TM} \omega\ d\mu \tag{2.16}$$

for any closed 1-form ω on M, where $[\omega] \in H^1(M,\mathbb{R})$ is the cohomology class of ω. Here we have used the identification([4]) $H_1(M,\mathbb{R}) \cong H^1(M,\mathbb{R})^*$ and equation (2.16) shows that $\rho(\mu)$ acts effectively on $H^1(M,\mathbb{R})$. Moreover since μ is holonomic the integral on (2.16) depends only on the cohomology class of ω; the class $\rho(\mu)$ is called the *homology* of μ or the *rotation* of μ by the analogy to the twist map theory. Using a finite basis $\{[\omega_1],...,[\omega_n]\}$ for $H^1(M,\mathbb{R})$ and the topology of the set of holonomic measure we get the following

Lemma 2.6.1. *The map $\rho : \overline{C(M)} \longrightarrow H_1(M,\mathbb{R})$ is continuous.*

Given a differentiable flow φ_t on a compact manifold N and a φ_t-invariant probability measure μ, the *Schwartzman's asymptotic cycle* of μ is defined to be the homology class $A(\mu) \in H_1(N,\mathbb{R}) \cong H^1(N,\mathbb{R})^*$ such that

$$< A(\mu), [\omega] > \ = \ \int_N \omega(X)\ d\mu$$

for any closed 1-form ω, where $[\omega]$ is the cohomology class of ω and X is the vector field whose flow is φ_t. This integral depends only on the cohomology class of ω because the integral of a coboundary by an invariant measure is zero; indeed if df is an exact 1-form, then

$$\begin{aligned}\int_N df(x)[v]\,d\mu \ &= \ \frac{1}{T}\int_0^T dt \int_N df(x)[v]\,d(\varphi_t)_*\mu = \frac{1}{T}\int_0^T dt\int_N df(x_t)[v_t]d\mu = \\ &= \ \frac{1}{T}\int_N d\mu \int_0^T df(x_t)[v_t]\,dt = \int_N \frac{f(x_T)-f(x_0)}{T}\,d\mu \ \overset{T\to+\infty}{\longrightarrow} \ 0\end{aligned}$$

because f is bounded; here we use the notation $(x_t, v_t) = \varphi_t(x_0, v_0)$. Moreover if μ is an ergodic probability measure and $x \in N$ is a *generic point*([5]) for μ, then

$$< A(\mu), [\omega] > \ = \ \lim_{T\to+\infty} \frac{1}{T}\int_0^T \omega(X(\varphi_t x))\,dt$$

and hence applying this to a basis $\{[\omega_1],...,[\omega_n]\}$ for $H^1(N,\mathbb{R})$ we get

$$A(\mu) \ = \ \lim_{T\to+\infty} \frac{1}{T}[\gamma_T * \delta_T] \in H_1(N,\mathbb{R})$$

where $\gamma_T(t) = \varphi_t(x)$, $t \in [0,T]$ and the curve δ_T is a unit speed geodesic from $\varphi_T(x)$ to x. In the case of a lagrangian flow the phase space $N = TM$ is not compact, but it has the same homotopy type

[4] Since M is compact $H_1(M,\mathbb{R})$ is a finite dimensional vector space and hence it is naturally isomorphic to its double dual $H^1(M,\mathbb{R})^*$.

[5] A point $x \in N$ is *generic* for an ergodic probability μ if $\lim_{T\to+\infty} \frac{1}{T}\int_0^T f(\varphi_t x)\,dt = \int_N f\ d\mu$ for all $f \in C^0(N,\mathbb{R})$.

as the configuration space M because M is a deformation retract of TM (contracting TM along the fiber to the zero section $M \times \{0\}$); moreover the ergodic components of an invariant measure of a lagrangian flow are contained in a unique energy level, which is a compact submanifold of TM. We see that the homology of an invariant probability and its asymptotic cycle coincide under the identification

$$H_1(TM, \mathbb{R}) \overset{\pi_*}{\cong} H_1(M, \mathbb{R})\,.$$

Lemma 2.6.2. *$\pi_*(A(\mu)) = \rho(\mu)$ for all $\mu \in \mathcal{M}(L)$.*

Proof. If ω is a closed 1-form on M then

$$(\pi^*\omega)(X(x,v)) = \omega[d\pi(X(x,v))] = \omega_x(v)$$

because the lagrangian vector field has the form $X(x,v) = (v, \cdot)$; hence

$$< \pi_* A(\mu), [\omega] > = < A(\mu), \pi^*[\omega] > = \int_{TM} (\pi^*\omega)(X)\; d\mu \;= \int_{TM} \omega\; d\mu \;= < \rho(\mu), [\omega] > \;.$$

□

Lemma 2.6.3. *The map $\rho : \mathcal{M}(L) \longrightarrow H_1(M, \mathbb{R})$ is surjective.*

Proof. Let $h \in H_1(M, \mathbb{Z})$ be an integer homology class and let $\eta : [0,1] \longrightarrow M$ be a closed curve with homology class h. Moreover let $\gamma : [0,1] \longrightarrow M$ be a minimizer of the action of L among the set of absolutely continuous curves whose interval of definition is $[0,1]$ with the same homotopy class as η; then γ is is a periodic orbit for the lagrangian flow with period 1. The invariant measure μ_γ satisfies $\rho(\mu_\gamma) = h$; now observe that the map ρ is affine and $\mathcal{M}(L)$ is convex, hence $\rho(\mathcal{M}(L))$ is convex and, in particular, it contains the convex hull of $H_1(M, \mathbb{Z})$; thus $H_1(M, \mathbb{R}) \subseteq \rho(\mathcal{M}(L))$. □

Recall that the functional $\mathbb{A}_L : \mathcal{M}(L) \longrightarrow \mathbb{R}$ is lower semicontinuous; moreover the sets

$$\mathcal{M}(h) := \{\mu \in \mathcal{M}(L) \mid \rho(\mu) = h\}$$

are closed. Hence we can define the *Mather's beta function* $\beta : H_1(M, \mathbb{R}) \longrightarrow \mathbb{R}$ as

$$\beta(h) := \min_{\mu \in \mathcal{M}(h)} \mathbb{A}_L(\mu)\,;$$

we shall see later that the β-function is convex. First we define the *Mather's alpha function* as the convex dual of the β-function $\alpha := \beta^* : H^1(M, \mathbb{R}) \longrightarrow \mathbb{R}$, i.e.

$$\begin{aligned}
\alpha([\omega]) &= \max_{h \in H_1(M,\mathbb{R})} \Big[< [\omega], h > - \beta(h) \Big] = - \min_{\mu \in \mathcal{M}(L)} \Big[\mathbb{A}_L(\mu) - < [\omega], \rho(\mu) > \Big] = \\
&= - \min_{\mu \in \mathcal{M}(L)} \Big[\mathbb{A}_L(\mu) - \int_{TM} \omega\; d\mu \Big] = - \min_{\mu \in \mathcal{M}(L)} \mathbb{A}_{L-\omega}(\mu) \;=\; c(L - \omega)
\end{aligned}$$

and since $L - \omega$ is also a convex superlinear lagrangian the value $\alpha([\omega])$ is finite.

Theorem 2.6.4. *The α and β functions are convex and superlinear.*

Proof. We prove that the function β is convex; let $h_1, h_2 \in H_1(M, \mathbb{R})$ and $\mu_1, \mu_2 \in \mathcal{M}(L)$ such that

$$\rho(\mu_i) = h_i\,, \quad \mathbb{A}_L(\mu_i) = \beta(h_i) \qquad \text{for } i = 1, 2\,.$$

Then the probability $\nu := \lambda\mu_1 + (1-\lambda)\mu_2$ satisfies $\rho(\nu) = \lambda h_1 + (1-\lambda)h_2$ and hence

$$\beta(\lambda h_1 + (1-\lambda)h_2) \le \mathbb{A}_L(\lambda\mu_1 + (1-\lambda)\mu_2) = \lambda\beta(\mu_1) + (1-\lambda)\beta(\mu_2)\,.$$

We know that the map ρ is surjective, hence β is finite; thus by the proposition A.2.3 in the appendix the function α is superlinear. Moreover by the same proposition the functions α and β are both convex and finally β is superlinear because α is finite. □

Fixed a homology class $h \in H_1(M,\mathbb{R})$ and a cohomology class $\omega \in H^1(M,\mathbb{R})$ we can define the sets

$$\mathcal{M}_h(L) := \{\mu \in \mathcal{M}(L) \mid \rho(\mu) = h,\ \mathbb{A}_L(\mu) = \beta(h)\};$$

$$\mathcal{M}^\omega(L) := \{\mu \in \mathcal{M}(L) \mid \mathbb{A}_{L-\omega}(\mu) = -c(L-\omega)\};$$

Since the β-function has a supporting hyperplane at each homology class h, if $\omega \in \partial\beta(h)$, then $\mathcal{M}_h(L) \subseteq \mathcal{M}^\omega(L)$; conversely, since $\alpha^* = \beta$, we have $\mathcal{M}^\omega(L) \subseteq \mathcal{M}_h(L)$ if $h \in \partial\alpha(\omega)$. Thus we get

$$\bigcup_{h\in H_1(M,\mathbb{R})} \mathcal{M}_h(L) \ = \ \bigcup_{\omega\in H^1(M,\mathbb{R})} \mathcal{M}^\omega(L)\,;$$

we call these measures *Mather minimizing measures* and the set

$$\mathcal{M} := \mathcal{M}^0(L) = \{\mu \in \mathcal{M}(L) \mid \mathbb{A}_L(\mu) = -c(L)\}$$

the *Mather set.* Finally we can define the *strict critical value* as

$$c_0(L) := \min_{\omega\in H^1(M,\mathbb{R})} c(L-\omega) = \min_{\omega\in H^1(M,\mathbb{R})} \alpha(\omega) = -\beta(0)\,.$$

Remark. *Using the characterization of minimizing measures 2.5.3 and corollary 2.3.2 we get the following chain of inclusions*([6])

$$\mathcal{M} \ \subseteq \ \mathcal{A} \ \subseteq \ \widetilde{\mathcal{N}} \ \subseteq \ \acute{\mathcal{E}}\,,$$

where $\acute{\mathcal{E}}$ is the energy level $\acute{\mathcal{E}} = \{E \equiv c(L)\}$. All these inclusions can be made proper by constructing examples of embedded flows (cfr. Contreras [6], chapter 1) and adding a properly chosen potential $U(x)$. Moreover corollary 2.3.2 implies that the strict critical value is the lowest energy level which supports Mather minimizing measures and, since by definition $c_0(L) = -\beta(0)$, these minimal energy Mather minimizing measures have trivial homology.

2.7 Coverings: the critical values $c_u(L)$, $c_a(L)$

In the previous section we discussed about compact manifolds, but there some important non compact cases as for example coverings. Particularly interesting are the *abelian cover* $\widehat{M}$ of a manifold M, whose fundamental group is the kernel of the Hurewicz homomorphism

$$h : \pi_1(M) \longrightarrow H_1(M,\mathbb{Z})$$

and whose deck transformation group is $H_1(M,\mathbb{Z})$, and the *universal cover* $\widetilde{M}$([7]). Observe that a closed curve in $\widehat{M}$ projects to a closed curve in M with trivial homology. If $M_1 \xrightarrow{p} M$ is a covering and $L : TM \longrightarrow \mathbb{R}$ is a lagrangian, denote by $L_1 := L \circ dp : TM_1 \longrightarrow \mathbb{R}$ the lifted lagrangian.

Lemma 2.7.1. *If $M_1 \xrightarrow{p} M$ is a covering then $c(L_1) \le c(L)$.*

Proof. It is obvious because closed curves on M_1 project to closed curves on M. □

Proposition 2.7.2. *If $M_1 \xrightarrow{p} M$ is a finite covering then $c(L_1) = c(L)$.*

Proof. We know that $c(L_1) \le c(L)$; by contradiction suppose that the strict inequality holds and let $k \in \mathbb{R}$ such that $c(L_1) < k < c(L)$. Hence there exists a closed curve γ in M with negative $(L+k)$-action and since M_1 is a finite covering of M some iterate of γ lifts to a closed curve on M_1 with negative (L_1+k)-action and this yields to a contradiction with the hypothesis $c(L_1) < k$. □

Definition 2.7.3. *Given a manifold M and a lagrangian $L : TM \longrightarrow \mathbb{R}$ we define the* ***critical value of the universal cover*** *$c_u(L)$ as the Mañé's critical value of the lifted lagrangian to the universal cover and analogously the* ***critical value of the abelian cover*** *$c_a(L)$ as the Mañé's critical value of the lifted lagrangian to the abelian cover, i.e.*

[6]The typographical relationship was observed by Albert Fathi.

[7]When $\pi_1(M)$ is abelian these two coverings coincide.

$$c_u(L) := c(\mathbb{L}_{\widetilde{M}})\,, \qquad c_a(L) := c(\mathbb{L}_{\widehat{M}})$$

where $\mathbb{L}_{\widetilde{M}} = L \circ dp : T\widetilde{M} \longrightarrow \mathbb{R}$ *and* $\mathbb{L}_{\widehat{M}} = L \circ dp : T\widehat{M} \longrightarrow \mathbb{R}$.

Proposition 2.7.4. *If L is a lagrangian then $c_0(L) = c_a(L)$.*

Proof. Let ω be a closed form on M; since $H^1(\widehat{M}, \mathbb{R}) = \{0\}$ the lift $\hat{\omega}$ of ω to $\widehat{M}$ is exact, then

$$c_a(L) := c(\mathbb{L}_{\widehat{M}}) = c(\mathbb{L}_{\widehat{M}} - \hat{\omega}) \leq c(L - \omega)$$

where the equality follows from the fact that adding an exact form to a lagrangian does not change the critical value while the inequality follows from lemma 2.7.1 and hence we get

$$c_a(L) \leq \min_{\omega \in H^1(M,\mathbb{R})} c(L - \omega) = c_0(L)\,.$$

Conversely we know that

$$\begin{aligned} -c_a(L) \;=\; -c(\mathbb{L}_{\widehat{M}}) \;&=\; \inf\left\{\mathbb{A}_{\mathbb{L}_{\widehat{M}}}(\mu) \;\middle|\; \mu \in \overline{C(\widehat{M})}\right\} = \\ &=\; \inf\left\{\mathbb{A}_{\mathbb{L}_{\widehat{M}}}(\mu_\gamma) \;\middle|\; \mu_\gamma \in C(\widehat{M})\right\} = \\ &=\; \inf\left\{\mathbb{A}_L(\mu_\gamma) \;\middle|\; \mu_\gamma \in C(M),\; \rho(\mu_\gamma) = 0\right\} \end{aligned}$$

because a closed curve γ on M has homology $[\gamma] = 0$ if and only if it has a closed lift to $\widehat{M}$; then by theorem 2.4.2, together with the continuity of the function ρ, it follows that

$$\begin{aligned} -c_a(L) \;&=\; \inf\left\{\mathbb{A}_L(\mu) \;\middle|\; \mu \in \overline{C(M)},\; \rho(\mu) = 0\right\} \leq \\ &\leq\; \min\left\{\mathbb{A}_L(\mu) \;\middle|\; \mu \in \mathcal{M}(L),\; \rho(\mu) = 0\right\} \;=\; \beta(0) \;= -c_0(L) \end{aligned}$$

and this concludes the proof. $\square$

Given a lagrangian L the set $\Sigma^+(L)$ is defined by

$$\Sigma^+(L) := \{w \in TM \mid x_w : [0, +\infty) \longrightarrow M \;\text{ is semistatic}\}\,;$$

it is possible to prove the following covering properties (see [6] for the proof).

Theorem 2.7.5. *The projection $\pi\big|_{\Sigma^+(L)} : \Sigma^+(L) \longrightarrow M$ is surjective.*

Altough in general the projection in theorem 2.7.5 is not injective, corollary 2.3.2 implies $E(\Sigma^+(L)) \equiv c(L)$ and then using the equivalent definition of e_0 given by

$$e_0 \;=\; \min\,\{k \in \mathbb{R} \mid \pi : E^{-1}(k) \longrightarrow M \;\text{ is surjective}\}$$

we get the following chain of real numbers

$$e_0(L) \leq c_u(L) \leq c_a(L) = c_0(L) \leq c(L)\,.$$

When $c_u(L) < c_0(L)$ the method explained in equation (2.14) allows to obtain minimizing measures which are not Mather minimizing; however, for symmetric lagrangians, we have $c_u(L) = c_0(L) = c_a(L) = c(L)$. Mañé in [9] gives an example in which $e_0 < c_a(L) < c(L)$, while G. and M. Paternain in [10] give an example in which $c_u(L) < c_a(L)$.

Chapter 3

The Hamiltonian viewpoint

3.1 Lagrangian graphs

In chapter 1 we have pointed out that near the lagrangian approach to the study of dynamical systems there is another equivalent approach, called Hamiltonian formalism; this approach sometimes can be more fruitful and, in particular, allows to use general results as for instance those in symplectic topology([1]) and at the same time to develop new tools coming from the study of *Lagrangian graphs*, *Floer homology* and *weak KAM theory*. In this chapter we are going to discuss only about the first argument; the reader can refer to [26] for a survey on Floer homology and to [7] or to [27] for a good introduction to the third argument, although the fundamental reference is the still unpublished book of Albert Fathi *Weak KAM theorem in Lagrangian dynamics*. As usual denote by ω the canonical symplectic form on T^*M.

Definition 3.1.1. *Let M be a n-manifold; a subspace V of $T_\xi T^*M$ is called* ***isotropic***([2]) *if $\omega|_V \equiv 0$, i.e. if $\omega(x, y) = 0$ for all $x, y \in V$. If V is an isotropic subspace and $\dim V = n$ we say that V is a lagrangian subspace of $T_\xi T^*M$. A submanifold $W \subseteq T^*M$ is called* ***lagrangian*** *if at each point $\xi \in W$ the tangent space $T_\xi W$ is a lagrangian subspace of $T_\xi T^*M$; in particular $\dim W = \dim M$.*

Theorem 3.1.2 (Hamilton-Jacobi)**.** *If the hamiltonian H is constant on a lagrangian submanifold N then N is invariant under the hamiltonian flow.*

Proof. We only have to show that the hamiltonian vector field X_H is tangent to N; since H is constant on N we have that $dH|_{TN} \equiv 0$ and since by definition

$$\omega(X_H, \cdot) = -dH$$

we get $\omega(X, \xi) = -dH(\xi) = 0$ for all $\xi \in TN$. But we know that the tangent spaces to N are lagrangian, i.e. they are maximal isotropic subspaces, and this implies that $X \in TN$. □

Some distinguished and important n-dimensional submanifolds on T^*M are the graph submanifolds, which are of the form

$$G_\eta = \{(x, \eta_x) \mid x \in M\} \subseteq T^*M$$

where η is a 1-form on M; a *lagrangian graph* is a lagrangian graph submanifold. By working in local coordinates one can prove the following

Lemma 3.1.3. *G_η is a lagrangian graph if and only if η is closed.*

Proof. Let $\{q_1, ..., q_n\}$ be a system of local coordinates of M; this induces a local coordinate system $\{q_1, ..., q_n, p_1, ..., p_n\}$ of T^*M and in these coordinates the 1-form η is given by

[1] Recall that the hamiltonian is defined on the cotangent bundle T^*M which is naturally equipped with a symplectic structure.

[2] Since ω is non degenerate the isotropic subspaces have dimension $\leq n$, half of the dimension of T^*M.

$$\eta \;=\; \sum_{i=1}^{n} p_i(q)\, dq_i\,.$$

Hence a basis for TG_η is

$$E_i := \left(\frac{\partial}{\partial q_i}\,,\, \sum_{k=1}^{n} \frac{\partial p_k}{\partial q_i} \cdot \frac{\partial}{\partial p_k} \right) \qquad i = 1, ..., n\,;$$

A simple computation shows that

$$\begin{aligned}
\omega(E_i, E_j) &= (dp \wedge dq) \left[\left(\frac{\partial}{\partial q_i}\,,\, \sum_{k=1}^{n} \frac{\partial p_k}{\partial q_i} \cdot \frac{\partial}{\partial p_k} \right)\,,\, \left(\frac{\partial}{\partial q_j}\,,\, \sum_{k=1}^{n} \frac{\partial p_k}{\partial q_j} \cdot \frac{\partial}{\partial p_k} \right) \right] = \\
&= \left(\sum_{h=1}^{n} dp_h \wedge dq_h \right) \left[\left(\frac{\partial}{\partial q_i}\,,\, \sum_{k=1}^{n} \frac{\partial p_k}{\partial q_i} \cdot \frac{\partial}{\partial p_k} \right)\,,\, \left(\frac{\partial}{\partial q_j}\,,\, \sum_{k=1}^{n} \frac{\partial p_k}{\partial q_j} \cdot \frac{\partial}{\partial p_k} \right) \right] = \\
&= \frac{\partial p_i}{\partial q_j} - \frac{\partial p_j}{\partial q_i}
\end{aligned}$$

and at the same time

$$d\eta \;=\; \sum_{i,j=1}^{n} \frac{\partial p_i}{\partial q_j}(q)\, dq_j \wedge dq_i \;=\; \sum_{i<j} \left(\frac{\partial p_i}{\partial q_j} - \frac{\partial p_j}{\partial q_i} \right) dq_j \wedge dq_i\,;$$

therefore $d\eta = 0$ if and only if $\omega\big|_{TG_\eta} = 0$. □

We can associate a cohomology class $[\eta] \in H^1(M, \mathbb{R})$ to each lagrangian graph G_η; lagrangian graphs with zero cohomology class are the graphs of exact 1-forms G_{df} where $f : M \longrightarrow \mathbb{R}$ is a smooth function and they are called *exact lagrangian graphs*. If $H : T^*M \longrightarrow \mathbb{R}$ is a hamiltonian then the *Hamilton-Jacobi equation* associated to H is

$$H(x, du(x)) = k\,, \qquad u : M \longrightarrow \mathbb{R}\,; \tag{H-J}$$

hence a smooth solution $u : M \longrightarrow \mathbb{R}$ of the Hamilton-Jacobi equation corresponds to an exact invariant lagrangian graph. More generally a function $u : M \longrightarrow \mathbb{R}$ is called a *subsolution* of the Hamilton-Jacobi equation if for all $x \in M$ it satisfies

$$H(x, du(x)) \leq k\,.$$

Theorem 3.1.4. *If M is any covering of a closed manifold then*

$$\begin{aligned}
c(L) &= \inf_{u \in C^\infty(M)} \sup_{x \in M} H(x, du(x)) = \\
&= \inf\, \{k \in \mathbb{R} \mid \exists u \in C^\infty(M, \mathbb{R}) \text{ such that } H(x, du(x)) < k\}\,.
\end{aligned}$$

This theorem could be restated by saying that $c(L)$ is the infimum of the values of $k \in \mathbb{R}$ for which the sublevel $\{H < k\} = H^{-1}(-\infty, k)$ contains an exact lagrangian graph; this is a very geometric way to describe it. In the theorem above we consider only exact lagrangian graphs; if we consider lagrangian graphs with other cohomology classes, we obtain Mather's α-function.

Corollary 3.1.5. *If M is a compact manifold then*

$$\alpha(k) \;=\; \inf_{[\omega]=k} \sup_{x \in M} H(x, \omega(x))\,.$$

Proof. Let us fix a closed form ω_0 such that $[\omega_0] = k$; as we have pointed out in section 2.6 we have that $\alpha(k) = c(L - \omega_0)$. Hence, it suffices to show that

$$c(L-\omega_0) \;=\; \inf_{[\omega]=k} \sup_{x\in M} H(x,\omega(x))\,;$$

this follows easily from the theorem above by noticing that the hamiltonian associated with $\widehat{L} = L - \omega_0$ is $H(x, p + \omega_0(x))$, as one can see applying the Fenchel transform

$$\begin{aligned}\widehat{H}(x,p) \;&=\; \max_{v\in T_xM}\Big[<p,v>_x - \widehat{L}(x,v)\Big] \;=\; \max_{v\in T_xM}\Big[<p,v>_x - L(x,v) + \omega_0(x,v)\Big] \;=\\ &=\; \max_{v\in T_xM}\Big[<p+\omega_0,v>_x - L(x,v)\Big] \;=\; H(x,p+\omega_0(x))\end{aligned}$$

and that all the closed forms in the class k are given by $\omega_0 + df$ where f is a smooth function. □

In particular the critical value of the abelian cover $c_0(L)$ is the infimum of the energy levels which contain a lagrangian graph of any cohomology class in its interior. We do not give all the details of the proof of theorem 3.1.4 (for further details see [6]); however, before starting the proof, we have to talk for a while about *dominated functions* and *weak solutions* of the Hamilton-Jacobi equation.

3.2 Dominated functions

Definition 3.2.1. *We say that a function $u : M \longrightarrow \mathbb{R}$ is* ***dominated*** *by $L + k$, $k \geq c(L)$ and we write $u \prec L + k$, if it satisfies*

$$u(y) - u(x) \leq \Phi_k(x,y) \qquad \forall\, x,y \in M\,.$$

Notice that, since $\Phi_k(y, y) \geq 0$ for all $y \in M$ whenever $k \geq c(L)$, we have

$$-\Phi_k(y,y) + \Phi_k(x,y) \leq \Phi_k(x,y) \qquad \forall\, x \in M$$

and hence the function $v(x) := -\Phi_k(x, y)$ is dominated by $L + k$ for any $y \in M$ while, since the triangle inequality, we have

$$\Phi_k(x,y) + \Phi_k(y,x) \geq \Phi_k(y,y) \quad\Longrightarrow\quad \Phi_k(y,y) - \Phi_k(y,x) \leq \Phi_k(x,y)$$

and hence the function $u(x) := \Phi_k(y, x)$ is dominated for any $y \in M$.

Lemma 3.2.2. *The following hold:*

1. *If $u \prec L+k$ then u is Lipschitz with the same Lipschitz constant as Φ_k; in particular a family of dominated functions is equicontinuous.*
2. *If $u \prec L + k$ then $H(x, du(x)) \leq k$ at any differentiability point x of u.*

Proof. To prove statement 1 it suffices to observe that, by the definition of a dominated function and proposition 2.1.1, u satisfies

$$u(y) - u(x) \;\leq\; \Phi_k(x,y) \;\leq\; c(k)\, d_M(x,y)$$

where $c(k) > 0$ is the Lipschitz constant of Φ_k and then we get the thesis by changing the role of x and y. Now let us prove item 2; we have that

$$u(y) - u(x) \;\leq\; \Phi_k(x,y) \;\leq\; \int_0^{T(\gamma)} \Big[L(\gamma(t),\dot\gamma(t)) + k\Big]\,dt$$

for all curves $\gamma \in C(x, y)$. This implies that

$$du(x)[v] \;\leq\; L(x,v) + k \quad\Longrightarrow\quad du(x)[v] - L(x,v) \;\leq\; k$$

for all $v \in T_xM$ when u is differentiable at $x \in M$; now since the Hamiltonian H is defined as the Fenchel transform of the Lagrangian L we get

$$H(x, du(x)) \;=\; \sup_{v \in T_xM} \Big[du(x)[v] - L(x,v)\Big] \;\leq\; k$$

and this completes the proof. □

Definition 3.2.3. *Given a dominated function $u \prec L+k$ we say that an absolutely continuous curve $\gamma : [a,b] \longrightarrow M$* ***realizes*** *$u$ if it satisfies*

$$u(\gamma(t)) - u(\gamma(s)) \;=\; \mathbb{A}_{L+k}\big(\gamma\big|_{[s,t]}\big) \qquad \forall\, a \leq s \leq t \leq b\,.$$

Notice that if γ is an absolutely continuous curve which realizes a dominated function u then

$$\mathbb{A}_{L+k}\big(\gamma\big|_{[a,b]}\big) \;=\; u(\gamma(b)) - u(\gamma(a)) \;\leq\; \Phi_k(\gamma(a), \gamma(b))$$

and hence γ is a global minimizer; in particular if $k = c(L)$ then γ is a semistatic orbit. The following proposition shows that we actually get a solution of the Hamilton-Jacobi equation if there are (semistatic) curves which realize a dominated function u.

Proposition 3.2.4. *Suppose that $u \prec L+k$, then:*

1. *If $\gamma : \,]-\epsilon, \epsilon[\, \longrightarrow M$ realizes u, then u is differentiable at $\gamma(0)$.*
2. *If $\gamma : \,]-\epsilon, 0\,] \longrightarrow M$ (or $\gamma : [0, \epsilon[\, \longrightarrow M$) realizes u and u is differentiable at $x = \gamma(0)$, then $du(x) = L_v(x, \dot\gamma(0))$ and $H(x, du(x)) = k$.*

Proof. Let $w \in T_xM$ and let $\eta_s(t)$ be a variation([3]) of γ fixing the endpoints $\gamma(-\epsilon), \gamma(\epsilon)$ and such that $\frac{\partial}{\partial s}\eta_s(0)\big|_{s=0} = w$; define now

$$\mathcal{A}(s) := \int_{-\epsilon}^{0} \Big[L\big(\eta_s(t), \frac{\partial}{\partial t}\eta_s(t)\big) + k\Big]\, dt\,.$$

Then integrating by parts and using the fact that γ satisfies the Euler-Lagrange equation we get

$$\begin{aligned}
\mathcal{A}'(0) \;&=\; \Big[L_v \frac{\partial}{\partial s}\eta_s(t)\Big|_{s=0}\Big]_{-\epsilon}^{0} + \int_{-\epsilon}^{0} \Big[L_x - \frac{d}{dt}L_v\Big] \frac{\partial}{\partial s}\eta_s(t)\Big|_{s=0} dt \;= \\
&=\; L_v\big(x, \dot\gamma(0)\big)\cdot w - L_v\left(\eta_s(\epsilon), \frac{\partial}{\partial t}\eta_s(t)\Big|_{t=-\epsilon}\right) \cdot \frac{\partial}{\partial s}\eta_s(-\epsilon)\Big|_{s=0} \;=\; L_v\big(x, \dot\gamma(0)\big)\cdot w
\end{aligned}$$

where the last equality follows from the fact that η_s leaves the endpoints fixed (indeed this implies $\frac{\partial}{\partial s}\eta_s(-\epsilon)\big|_{s=0} = 0$). At the same time using the facts that $u \prec L+k$ and that γ realizes u we get

$$\begin{aligned}
u(\eta_s(0)) - u(\eta_s(-\epsilon)) \;&\leq\; \mathbb{A}_{L+k}\big(\eta_s\big|_{[-\epsilon,0]}\big) \;=\; \mathcal{A}(s); \\
u(\gamma(0)) - u(\gamma(-\epsilon)) \;&=\; \mathbb{A}_{L+k}\big(\gamma\big|_{[-\epsilon,0]}\big) = \mathcal{A}(0);
\end{aligned}$$

and hence the following inequality

$$\frac{1}{s}\Big[u(\eta_s(0)) - u(x)\Big] \;=\; \frac{1}{s}\Big[u(\eta_s(0)) - u(\eta_s(-\epsilon)) + u(\gamma(-\epsilon)) - u(\gamma(0))\Big] \;\leq\; \frac{1}{s}\Big[\mathcal{A}(s) - \mathcal{A}(0)\Big]\,.$$

We have just proved that

$$\limsup_{s\to 0} \frac{1}{s}\Big[u(\eta_s(0)) - u(x)\Big] \;\leq\; \mathcal{A}'(0)\,; \tag{3.1}$$

[3] By definition a variation η_s of γ satisfies $\eta_0(t) = \gamma(t)$ for all $t \in\,]-\epsilon, \epsilon[$.

similarly one can define $\mathcal{B}(s) := \mathbb{A}_{L+k}\big(\eta_s\big|_{[0,\epsilon]}\big)$ and prove that

$$\limsup_{s\to 0} \frac{1}{s}\Big[u(x) - u(\eta_s(0))\Big] \;\le\; \mathcal{B}'(0) \;=\; -L_v\big(x,\dot\gamma(0)\big)\cdot w\,.$$

Thus we get

$$\liminf_{s\to 0} \frac{1}{s}\Big[u(\eta_s(0)) - u(x)\Big] \;\ge\; L_v\big(x,\dot\gamma(0)\big)\cdot w\,. \tag{3.2}$$

and combining inequalities (3.1) and (3.2) we obtain that u is differentiable at $x = \gamma(0)$. Now assume that $\gamma :]-\epsilon, 0] \longrightarrow M$ realizes u and that u is differentiable at $x = \gamma(0)$; the same argument in equation (3.1) shows that

$$du(x)[w] \;\le\; L_v\big(x,\dot\gamma(0)\big)\cdot w \qquad \forall\, w\in T_xM\,.$$

Applying this inequality to $-w$ and combining both inequalities we get that $du(x) = L_v\big(x,\dot\gamma(0)\big)$; hence, since $u \prec L{+}k$ and x is a differentiability point for u, by lemma 3.2.2 we have $H(x, du(x)) \le k$. Moreover, since γ realizes u, we have

$$u(\gamma(0)) - u(\gamma(t)) \;=\; \mathbb{A}_{L+k}\big(\gamma\big|_{[t,0]}\big) \;=\int_t^0 \Big[L\big(\gamma(s),\dot\gamma(s)\big) + k\Big]\,ds\,;$$

hence $du(x)[\dot\gamma(0)] = L\big(\gamma(0),\dot\gamma(0)\big) + k$ and the thesis follows. □

3.3 Weak solutions of the Hamilton-Jacobi equation

In this section we are going to prove theorem 3.1.4, altough we do not give all the details of the proof; this theorem turns to be an immediate consequence of lemmas 3.3.2 and 3.3.3 below. First of all we briefly talk about the existence of solutions of the Hamilton-Jacobi equation; more precisely we prove that for $k \ge c(L)$ there are always Lipschitz solutions of the Hamilton-Jacobi equation. An important result due to Fathi (cfr. [13]) states that, in the case of a compact manifold M, the only energy level which supports a differentiable solution is $k = c(L)$; instead when M is non-compact there may be differentiable solutions also in the case $k > c(L)$.

Proposition 3.3.1. *If $k \ge c(L)$ then for any $y \in M$ the function $u(x) := \Phi_k(y,x)$ satisfies $H(x, du(x)) = k$ for almost every $x \in M$.*

Proof. Since u is a Lipschitz function, by Rademacher's theorem (cfr. [14]), it is differentiable at Lebesgue-almost every point; moreover, since u is dominated, by proposition 3.2.4 it suffices to show that u is one-sided realized at every point. If $k > c(L)$ we know that for all $x \ne y \in M$ there exists a finite-time global minimizer $\gamma \in C_T(y,x)$ with

$$\mathbb{A}_{L+k}(\gamma) \;=\; \Phi_k(y,x)\,.$$

By the triangle inequality the function

$$\delta(t) :=\; \mathbb{A}_{L+k}\big(\gamma\big|_{[0,t]}\big) - \Phi_k(y,\gamma(t))$$

is increasing, indeed if $t \ge t'$ we get $-\Phi_k(y,\gamma(t)) + \Phi_k(y,\gamma(t')) \ge -\Phi_k(\gamma(t'),\gamma(t))$ and hence

$$\begin{aligned}
\delta(t) - \delta(t') \;&=\; \mathbb{A}_{L+k}\big(\gamma\big|_{[0,t]}\big) - \Phi_k(y,\gamma(t)) - \mathbb{A}_{L+k}\big(\gamma\big|_{[0,t']}\big) + \Phi_k(y,\gamma(t') \;= \\
&=\; \mathbb{A}_{L+k}\big(\gamma\big|_{[t',t]}\big) - \Phi_k(y,\gamma(t)) + \Phi_k(y,\gamma(t')) \;\ge \\
&\ge\; \mathbb{A}_{L+k}\big(\gamma\big|_{[t',t]}\big) - \Phi_k(\gamma(t'),\gamma(t)) \;\ge\; 0
\end{aligned}$$

but on the other hand we have $\delta \ge 0$ and $\delta(T) = 0$; hence $\delta(t) \equiv 0$ and this implies that γ backward-realizes u at x. In the case $k = c(L)$ we have that u may be realized by an infinite semistatic orbit; the reader can see Contreras [6] in order to get a complete proof of this fact. □

Lemma 3.3.2. *If there exists a C^1 function $u : M \longrightarrow \mathbb{R}$ such that $H(x, du(x)) \leq k$ then $k \geq c(L)$.*

Proof. Recall that by the Fenchel transform

$$H(x,p) \;=\; \max_{v \in T_p M} \Big[< p, v >_x - \, L(x,v) \Big] ;$$

since $H(x, du(x)) \leq k$ it follows that for all $(x, v) \in TM$

$$du(x)[v] - L(x,v) \;\leq\; k \, .$$

Therefore if $\gamma : [0, T] \longrightarrow M$ is any absolutely continuous curve with $T > 0$ we get

$$\int_0^T \Big[L(\gamma(t), \dot\gamma(t)) + k \Big] \, dt \;=\; \int_0^T \Big[L(\gamma(t), \dot\gamma(t)) + k - du(\gamma(t))[\dot\gamma(t)] \Big] \, dt \;\geq\; 0$$

and thus $k \geq c(L)$. □

Remark. *If $u \in C^1(M, \mathbb{R})$ is a subsolution of the Hamilton-Jacobi equation for $k = c(L)$, then the lagrangian L can be replaced by the lagrangian*

$$\mathbb{L}(x,v) := \; L(x,v) - du(x)[v] + c(L) \;\geq\; 0 \, .$$

The new lagrangian is positive, has the same minimizing measures as L and its α and β functions are translates of those of L; moreover the static set $\widehat\Sigma(\mathbb{L})$ is contained in the level set $\mathbb{L} = 0$.

Lemma 3.3.3. *Let M be a riemannian covering of a compact manifold and suppose that*

$$\sup_{|v| \leq k} \left\| \frac{\partial L}{\partial x}(x,v) \right\| \;<\; +\infty \, ; \tag{3.3}$$

if $f : M \longrightarrow \mathbb{R}$ is weakly differentiable and a subsolution for the Hamilton-Jacobi equation

$$H(x, df(x)) \;\leq\; k \qquad \textit{for a.e. } x \in M$$

then for all $\delta > 0$ there exists $u \in C^\infty(M, \mathbb{R})$ such that $H(x, du(x)) < k + \delta$ for all $x \in M$.

See [6] for the proof; observe that, if $p : M \longrightarrow N$ is a covering of a compact manifold N and the lagrangian L on M is lifted from a lagrangian l on N, i.e. $L = l \circ dp$, then condition (3.3) follows; this fact in general is no longer true, as the counterexample

$$L(x,v) = \frac{1}{2}|v|^2 + \sin(x^2)$$

on $\mathbb{R}$ shows. Lemma 3.3.3 implies that for any $k > c(L)$ there exists a smooth subsolution $f \in C^\infty(M, \mathbb{R})$ of the Hamilton-Jacobi equation. Such a smooth subsolution can be obtained by regularizing the Lipschitz-function $u(x) := \Phi_{c(L)}(q, x)$, where $q \in M$ is a fixed point; indeed by the triangle inequality we have that $u \prec L + k$ while by lemma 3.2.2 u satisfies $H(x, du(x)) \leq c(L)$ at any point $x \in M$ in which $u(x)$ is differentiable (that is, by Rademacher's theorem, at Lebesgue-almost every point). Combining lemma 3.3.2 and 3.3.3 we also get the following

Corollary 3.3.4. *There are no weakly differentiable subsolutions of (H-J) for $k < c(L)$.*

3.4 Finsler metrics

Let M be a manifold and let $L : TM \longrightarrow \mathbb{R}$ be any Lagrangian; in this section we prove that if $k > c(L)$ then the Euler-Lagrange flow on the energy level $E = k$ is a reparametrization of the geodesic flow on the unit tangent bundle of a Finsler metric. This allows to borrow theorems from Finsler geometry; recall that if $k > c(L)$ then there exists always a smooth subsolution $u \in C^\infty(M)$ of the Hamilton-Jacobi equation

$$H(x, du(x)) \;<\; k$$

as lemma 3.3.3 states. Such a smooth subsolution can be used to replace the Lagrangian L with

$$\mathbb{L}(x,v) := L(x,v) - du(x)[v]\,;$$

indeed these Lagrangians have the same energy function and equivalent variational principles. Hence they have the same Euler-Lagrange flow, minimizing orbits, and the same action functional on closed curves and invariant measures; their action potential are related by

$$\Psi_k(x,y) \;=\; \Phi_k(x,y) + u(y) - u(x)$$

for any $x, y \in M$. Furthermore $\mathbb{L}(x,v) + k > 0$ for any $(x,v) \in TM$.

Definition 3.4.1. *A* ***Finsler metric*** *on a manifold M is a function $F : TM \longrightarrow \mathbb{R}$ satisfying:*

1. *F is smooth on $TM \setminus M \times \{0\}$.*
2. *For any fixed $x \in M$, $F(x,v) \geq 0$ for each $v \in T_xM$ and $F(x,v) = 0$ if and only if $v = 0$.*
3. *For any fixed $x \in M$, $F(x, \lambda v) = \lambda F(x,v)$ for each $v \in T_xM$ and for each $\lambda \geq 0$ (but not necessarily for $\lambda < 0$).*
4. *For any fixed $x \in M$, for any $v \in T_xM$ the Hessian $F^{(2)}$ at v is positive definite.*

In other words a Finsler metric on M is a non necessarily simmetric norm on T_xM for each $x \in M$ which varies smoothly on M. Given a Finsler metric F and an absolutely continuous curve γ we can define its *Finsler length* as

$$l_F(\gamma) := \int F(\gamma, \dot\gamma)\,.$$

Since the Finsler metric is homogeneous of degree one, the definition does not depend on the parametrization of the curve; now define the *Finsler distance*

$$D_F(x,y) := \inf_{\gamma \in C(x,y)} l_F(\gamma)\,.$$

Observe that the function $L = F^2$ is a Lagrangian satisfying the properties of definition 1.1.1 except the smoothness in $M \times \{0\}$, but since we are going to work with energy fixed, this will not be a problem. Such a Lagrangian is called a *Finsler Lagrangian*; the Euler-Lagrange flow of $L = F^2$ is called the *geodesic flow* of the Finsler metric F. As it happens for the geodesic flow of a riemannian metric on M, such a geodesic flow depends on the energy only through a uniform change of speed; hence it suffices to study the flow on the unit (with respect to F) sphere in TM.

Theorem 3.4.2. *[28],[29] If $k > c(L)$ then the Euler-Lagrange flow on the energy level $E \equiv k$ is a reparametrization of the geodesic flow of a Finsler metric on its unit tangent bundle. Moreover if $u \in C^\infty(M, \mathbb{R})$ is such that $H(x, du(x)) < k$, the Finsler lagrangian F^2 can be taken to be*

$$F(x,v) \;=\; L(x,v) + k - du(x)[v]$$

on the energy level $E \equiv k$; in particular if $k > -\inf L(x,v)$ then one can choose $u \equiv 0$.

Proof. We have already pointed out that if $k > c(L)$ then there exists a smooth function $u \in C^\infty(M)$ that satisfies the Hamilton-Jacobi equation $H(x, du(x)) < k$; observe that

$$H(x,0) \;=\; \max_{v \in T_xM} \Big[< 0, v >_x - L(x,v) \Big] \;=\; - \inf_{v \in T_xM} L(x,v)$$

so that if $k > -\inf L(x,v)$ we can choose $u \equiv 0$. Moreover if $H(x, du(x)) < k$ then it must be that $k > c(L) > e_0$ and hence the zero section is contained in the sublevel $\{E < k\}$; at the same time we have $L - du + k > 0$ and this allows us to define a Finsler metric on TM by

$$F(x,v) \;=\; L(x,v) - du(x)[v] + k$$

on the energy level $E = k$ and extend it by homogeneity. Since $k > c(L) = c(L - du)$, by proposition 2.1.5, for any $x \neq y$ there exists a global minimizer on $C(x,y)$ for $L + k - du$ which has energy k;

by the homogeneity of F we can restrict the curves in the definition of D_F to those with energy k and hence

$$\Phi_k(x,y) \;=\; D_F(x,y) + u(y) - u(x) \tag{3.4}$$

for any $x, y \in M$. To show that the Euler-Lagrange flow on the energy level $E = k$ is a reparametrization of the geodesic flow on the unit tangent bundle of F we only need to prove that sufficiently small Euler-Lagrange solutions with energy k are geodesic of F. Let $\mathbb{L} := L - du$; since equality (3.4) implies that any $(\mathbb{L} + k)$-global minimizer is a geodesic for F it is enough to show that sufficiently small Euler-Lagrange solutions with energy k are global minimizers. Fix $x \in M$ and a small neighbourhood $\mathcal{N}(x)$ of x such that for all $y \in \mathcal{N}(x)$ there exists a unique Euler-Lagrange solution contained in $\mathcal{N}(x)$ with energy k and joining x to y. Let Ψ_k be the action potential of $\mathbb{L}$; then define

$$\epsilon := \; \inf\{\Psi_k(x,z) \mid z \notin \mathcal{N}(x)\} \;>\; 0$$

and

$$\mathcal{M}(x) := \left\{ y \in M \;\middle|\; \Psi_k(x,y) \le \frac{\epsilon}{2} \right\}.$$

Then $\mathcal{M}(x)$ is a neighbourhood of x and $\mathcal{M}(x) \subseteq \mathcal{N}(x)$; by the triangle inequality any $(\mathbb{L}+k)$-global minimizer joining x to $y \in \mathcal{M}(x)$ must be contained in $\mathcal{N}(x)$. At the same time such a minimizer effectively exists because of proposition 2.1.5; hence all the small Euler-Lagrange solutions contained in $\mathcal{N}(x)$ joining x to a point $y \in \mathcal{M}(x)$ are global minimizers. □

3.5 Magnetic flows

Let (M, g) be a riemannian manifold and let $\tau : T^*M \longrightarrow M$ be the canonical projection; consider a closed 2-form σ on M and, as usual, the canonical symplectic form ω_0 on T^*M, that in local coordinates is given by $\omega_0 = dp \wedge dx$. Then we can define a *twisted symplectic form* on T^*M by([4])

$$\omega_\sigma := \; \omega_0 + \tau^*\sigma$$

and consider the Hamiltonian flow on T^*M defined by the form ω_σ and by the Hamiltonian

$$H(x,p) \;=\; \frac{1}{2}\,\|p\|_x^2$$

where $\|\cdot\|_x$ denotes the dual norm on T^*M induced by the metric g. This flow models the motion of a particle on M moving under the magnetic field σ and it is called *magnetic flow*. The Hamiltonian vector field X_H associated to H is defined by the identity

$$\omega_\sigma(X_H, \cdot) \;=\; -dH[\,\cdot\,]\,. \tag{3.5}$$

Let $u \in TM$ and let u_x, u_p, X_x, X_p be respectively the x,p-components of u, X_H; then

$$\begin{aligned}
\omega_\sigma(X_H, u) \;&=\; \omega_0(X_H,u) + \tau^*\sigma(X_H,u) \;=\; X_p \cdot u_x - X_x \cdot u_p + \sigma(X_x,u_x) \;= \\
&=\; X_p \cdot u_x - X_x \cdot u_p + \frac{1}{2}\sum \sigma_{ij}\, dx_j \wedge dx_j\,(X_x,u_x) \;= \\
&=\; X_p \cdot u_x - X_x \cdot u_p + \frac{1}{2}\sum \sigma_{ij}\,[X_{x_i}u_{x_j} - X_{x_j}u_{x_i}] \;= \\
&=\; X_p \cdot u_x - X_x \cdot u_p - \sum \sigma_{ij}\, X_{x_j}u_{x_i}
\end{aligned}$$

and at the same time $-dH[u] = -H_x \cdot u_x - H_p \cdot u_p$. Since (3.5) is true for any possible choice of u we get that the Hamiltonian vectorfield X_H is defined by the system

[4] Observe that ω_σ is effectively a symplectic form, since it is obviously closed and it can be easily checked that it is non degenerate.

$$\begin{cases} X_x = H_p \,; \\ \\ X_p = -H_x + \sum_{i,j} \sigma_{ij}(x)\, H_{p_j} \,; \end{cases}$$

so that the Hamiltonian system is

$$\begin{cases} \dot{x}_i = H_{p_i} \,; \\ \\ \dot{p}_i = -H_{x_i} + \sum_{j} \sigma_{ij}(x)\, H_{p_j} \,; \end{cases} \tag{3.6}$$

This is in a certain sense the easiest way to say what a magnetic flow is; in fact, a magnetic flow can be viewed also as a Hamiltonian flow on the tangent bundle TM. This approach can be sometimes more fruitful, since it allows, with some further conditions, to see a magnetic flow as the Euler-Lagrange flow of a suitable Lagrangian. Thus let (M, g) be a smooth Riemannian manifold and let $\pi : TM \longrightarrow M$ be the canonical projection; let Ω_0 be the symplectic form on TM obtained by pulling back the canonical symplectic form ω_0 on T^*M through the riemannian metric. If σ is a closed 2-form on M we can define a new symplectic form

$$\Omega_\sigma := \Omega_0 + \pi^*\sigma \;;$$

the *magnetic flow* of the pair (g, σ) is the Hamiltonian flow of the function

$$E(x, v) \;=\; \frac{1}{2}\|v\|_x^2 \tag{3.7}$$

with respect to Ω_σ. It models the motion of a particle of unit mass and charge under the effect of a magnetic field, whose Lorentz force $Y : TM \longrightarrow TM$ is the bundle map defined by

$$\sigma_x(u, v) \;= < Y_x(u), v >_x$$

for all $x \in M$ and for all $u, v \in T_xM$. In other words a curve $t \longmapsto (\gamma(t), \dot{\gamma}(t))$ is an orbit of the Hamiltonian flow of the pair (g, σ) if and only if

$$\frac{D}{dt}\dot{\gamma} \;=\; Y_\gamma(\dot{\gamma})\,. \tag{3.8}$$

A curve that satisfies equation (5.1) is called a *magnetic geodesic*; observe that the magnetic flow of the pair $(g, 0)$ is the geodesic flow of (M, g). Magnetic flows were first considered by V.I. Arnold in [30] and by D.V. Anosov and Y.G. Sinai in [31]; recent work on this flows has uncovered several remarkable properties, see for instance [23]. If σ is also exact, i.e. if there exists a 1-form ϑ such that $d\vartheta = \sigma$, then we can consider the Lagrangian $L : TM \longrightarrow \mathbb{R}$

$$L(x, v) := \;\frac{1}{2}\|v\|_x^2 + \vartheta_x(v)\,. \tag{3.9}$$

The extremals of L, i.e. the solutions of the Euler-Lagrange equation associated to L are the magnetic geodesics and the energy function is exactly the function (3.7). The problem is that the form σ in general is not exact; this will be, for instance, the case we are going to face in section 5.6. In this case we can not write anymore the Lagrangian in (3.9); however we can consider the universal covering (or more generally any arbitrary covering of M)

$$p : \widetilde{M} \longrightarrow M$$

of M and the form $\widetilde{\sigma}$ obtained by lifting σ to $\widetilde{M}$. If $\widetilde{\sigma}$ is exact, i.e. if there exists a 1-form ϑ on $\widetilde{M}$ such that $d\vartheta = \widetilde{\sigma}$, we can proceed in an analogous way as above and define a Lagrangian $L : T\widetilde{M} \longrightarrow \mathbb{R}$ by the formula (3.9). Now it turns to be that the Euler-Lagrange orbits of L coincide with the lift of the magnetic geodesics and the energy function of L is equal to the lift of E.

The magnetic flow shares with the geodesic flow the property that the level sets of E are invariant; there is, however, a significant difference. Indeed the geodesic flow is the same for all energy levels up to a uniform change of speed; on the other hand, for the magnetic flow the behaviour depends in an essential way on the energy and changes drastically crossing the *Mañé's critical value*

$$c(L) := \inf \left\{ k \in \mathbb{R} \mid \mathbb{A}_{L+k} \geq 0 \text{ for any a.c. closed curve } \gamma \text{ defined on any closed interval} \right\}.$$

In the case $\tilde{\sigma}$ exact, the magnetic flow on $T\widetilde{M}$, as any other Lagrangian flow, can be viewed as the Hamiltonian flow defined by the canonical symplectic form on $T^*\widetilde{M}$ and the Hamiltonian

$$H(x,p) \;=\; \frac{1}{2}\|p - \vartheta_x\|_x^2 \,.$$

The Legendre transform $\mathcal{L} : T\widetilde{M} \longrightarrow T^*\widetilde{M}$ carries orbits of the Lagrangian flow for L into orbits of the Hamiltonian flow defined by H and the canonical symplectic form. The *critical value of the pair* (g, σ) is defined as

$$c(g,\sigma) := \inf_{u \in C^\infty(\widetilde{M})} \sup_{x \in \widetilde{M}} H(x, d_x u) \quad = \inf_{u \in C^\infty(\widetilde{M})} \sup_{x \in \widetilde{M}} \frac{1}{2}\|d_x u - \vartheta_x\|_x^2 \,; \tag{3.10}$$

as u ranges among $C^\infty(\widetilde{M}, \mathbb{R})$ the form $\vartheta - du$ ranges over all primitives of $\widetilde{\sigma}$, because any two primitives differ by a closed 1-form which must be exact since the universal cover is simply connected. It is clear that $c(g,\sigma) \geq 0$; conversely, it is possible to prove that if σ is non trivial then $c(g,\sigma) > 0$; moreover it follows from theorem 3.1.4 that $c(L) = c(g,\sigma)$, even if all primitives of $\tilde{\sigma}$ are unbounded.

Chapter 4

The Horocycle flows

4.1 The hyperbolic plane $\mathbb{H}$

In this section we introduce the hyperbolic plane, that is the subset $\mathbb{H} = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ of the real plane $\mathbb{R}^2$ endowed with the Riemannian metric

$$g = \frac{1}{y^2}(dx \otimes dx + dy \otimes dy), \tag{4.1}$$

and briefly recall its properties (we refer to the appendix for a more detailed discussion). We often represent $\mathbb{H}$ also as a subset of the complex plane and indicate with $z = x + iy$ the standard coordinate on $\mathbb{C}$; these two characterization are equivalent and we will use both indistinctly. As one probably already know the hyperbolic plane is a complete riemannian surface and its geodesics are the vertical lines and the half-circles orthogonal to the real axis.

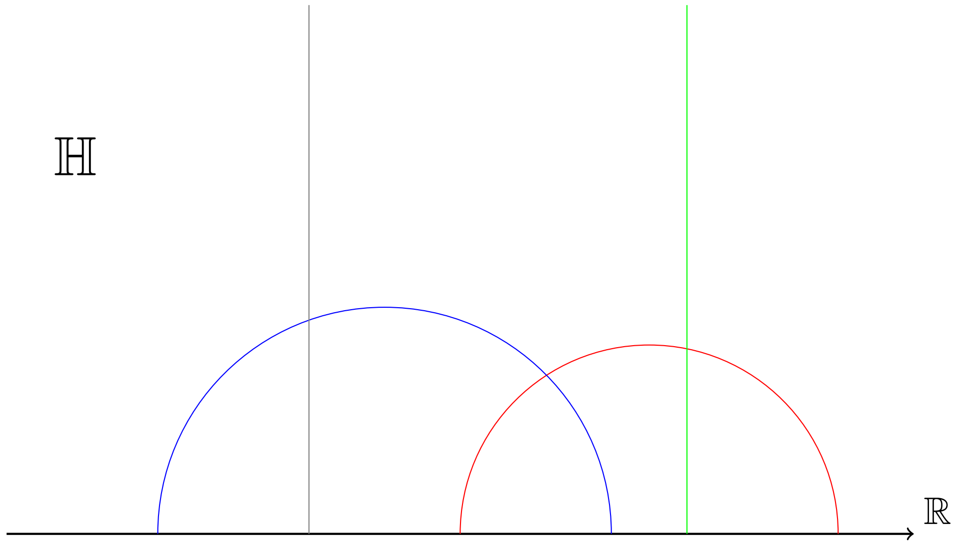


Figure B.1 *The geodesics of the hyperbolic plane are vertical lines or half-circles hortogonal to* $\mathbb{R}$.

Moreover $PSL(2, \mathbb{R})$ acts on $\mathbb{H}$ as a group of isometries through the map

$$\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] \longmapsto f(z) = \frac{az + b}{cz + d};$$

conversely one can prove that any isometry of $\mathbb{H}$ is of the form $z \longmapsto A(z)$ or $z \longmapsto A(\bar{z})$ for a suitable $A \in PSL(2, \mathbb{R})$. The action of $PSL(2, \mathbb{R})$ can be lifted to an action on $T\mathbb{H}$ by defining

$$A^*(z, v) := (A(z), dA(z)[v])$$

(here A acts on $\mathbb{H}$ and its action is not linear, so the differential is not A itself). Moreover since any elements of $PSL(2,\mathbb{R})$ acts as an isometry, it makes sense to restrict the action above at the unit tangent bundle $T_1\mathbb{H}$ and it turns to be that this action is transitive; hence we can map any unit tangent vector to any other using an element of $PSL(2,\mathbb{R})$. The isometries of $\mathbb{H}$ can also be classified with respect to their number of fixed points, in the sense that for any isometry $\varphi \neq id$ of $\mathbb{H}$ exactly one of the following holds:

- φ has a single fixed point and this lies on $\partial\mathbb{H}$, in this case we say that φ is *parabolic*.
- φ has two fixed points and these both lie on $\partial\mathbb{H}$, in this case we say that φ is *hyperbolic*.
- φ has a single fixed point in $\mathbb{H}$ and none in $\partial\mathbb{H}$, in this case we say that φ is *elliptic*.

Therefore if $\Gamma < PSL(2,\mathbb{R})$ is a group of hyperbolic and parabolic isometries acting properly discontinuously on $\mathbb{H}$ we get that $\pi : \mathbb{H} \longrightarrow \Gamma \setminus \mathbb{H}$ is a covering and $\Gamma \setminus \mathbb{H}$ is a complete hyperbolic (i.e. locally isometric to $\mathbb{H}$) surface. Conversely it is possible to prove that any complete iperbolic surface is obtained from the hyperbolic plane quotienting out by a group of parabolic and hyperbolic isometries acting properly discontinuously on $\mathbb{H}$ and that the quotient $\Gamma \setminus \mathbb{H}$ is compact if and only if Γ does not contain parabolic elements.

4.2 The geodesic flow

If $z \in \mathbb{H}$ is a point and $v \in T_z\mathbb{H}$ is a unit tangent vector, that is $\|v\|_z = 1$, then v determines uniquely a geodesic $\gamma_v : \mathbb{R} \longrightarrow \mathbb{H}$ such that $\gamma_v(0) = z$ and $\dot{\gamma}_v(0) = v$; such γ_v is either a (Euclidean) semicircle orthogonal to the real axis or a vertical line. The geodesics may often cross in $\mathbb{H}$, so to define a flow we must work on the unit tangent bundle $T_1\mathbb{H}$ by considering a geodesic as the pair $(\gamma_v, \dot{\gamma}_v)$; this pair is an orbit of the geodesic flow

$$g_t : T_1\mathbb{H} \longrightarrow T_1\mathbb{H}\,, \quad g_t(z,v) \;=\; (\gamma_v(t), \dot{\gamma}_v(t))\,.$$

As already observed in the previous section the action of $PSL(2,\mathbb{R})$ on $\mathbb{H}$ lifts to a transitive action on $T_1\mathbb{H}$; hence if we denote by $i_i \in T_i\mathbb{H}$ the unit vertical tangent vector at the point $i \in \mathbb{H}$ we can identify the unit tangent bundle $T_1\mathbb{H}$ with the set

$$\{A^*(i, i_i) \mid A \in PSL(2,\mathbb{R})\} \;\cong\; PSL(2,\mathbb{R})\,.$$

We want to see how the geodesic flow g_t is expressed under the identification $T_1\mathbb{H} \cong PSL(2,\mathbb{R})$; the geodesic γ_{i_i} is the imaginary axis parametrized by $\gamma_{i_i}(t) = ie^t$ and since the isometry $g(z) = e^t z$, whose associated matrix is

$$A_t \;= \begin{pmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{-\frac{t}{2}} \end{pmatrix}\,,$$

preserves the imaginary axis we have that $g_t(I) = A_t$. Observe that the geodesic flow commutes with the action of $PSL(2,\mathbb{R})$ on $T_1\mathbb{H}$, in the sense that the diagram

$$\begin{tikzcd} T_1\mathbb{H} \arrow[r, "A^*"] \arrow[d, "g_t"'] & T_1\mathbb{H} \arrow[d, "g_t"] \\ T_1\mathbb{H} \arrow[r, "A^*"'] & T_1\mathbb{H} \end{tikzcd}$$

is commutative for any $A \in PSL(2,\mathbb{R})$, where A^* is the lift of A to $T_1\mathbb{H}$; therefore if $B \in PSL(2,\mathbb{R})$ then the g_t-orbit is $g_t(B) = BA_t$. We have just proved that the geodesic flow is given by

$$g_t : PSL(2,\mathbb{R}) \longrightarrow PSL(2,\mathbb{R})\,, \quad g_t(B) = B\begin{pmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{-\frac{t}{2}} \end{pmatrix}\,.$$

Any complete hyperbolic surface is of the form $M = \Gamma \setminus \mathbb{H}$ where $\Gamma < PSL(2,\mathbb{R})$ is a group of hyperbolic and parabolic isometries acting properly discontinuously on $\mathbb{H}$; since the geodesics are invariant under Γ, quotienting out by the action of Γ we get the geodesics γ_v for any $v \in T_1M$. Moreover since the geodesic flow on $T_1\mathbb{H}$ commutes with the action of Γ, it projects to the geodesic flow g_t on T_1M; the unit tangent bundle T_1M can be seen as

$$\Gamma \setminus PSL(2,\mathbb{R}) = \{\Gamma B \mid B \in PSL(2,\mathbb{R})\}$$

and under this identification the geodesic flow is given by

$$g_t : T_1M \longrightarrow T_1M\,, \quad g_t(\Gamma B) \;=\; \Gamma B \begin{pmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{-\frac{t}{2}} \end{pmatrix}.$$

We end this section computing the arclength parametrisation of the geodesics of $\mathbb{H}$; we get this parametrisation by considering geodesics ending in the same point (the point at infinity) and then changing the ending point by an isometry. It is easy to see that if $v = i_{b+i} \in T_{b+i}\mathbb{H}$ then

$$\gamma_v^b(t) \;=\; b + ie^t \;=\; \big(b, e^t\big) \tag{4.2}$$

is the arclength parametrization of the vertical geodesic through the point $(b,1)$; the curves γ_v^b, $b \in \mathbb{R}$, parametrize the geodesics ending at ∞. Therefore we get the arclength parametrization of the geodesics ending at the origin by composing $\gamma_v^b(t)$ with a suitable isometry of $\mathbb{H}$ that maps the point at infinity into the origin, namely

$$T(x,y) \;=\; T(z) \;=\; \frac{-1}{z} \;=\; \frac{-1}{x+iy} \;=\; \frac{1}{x^2+y^2}\big(-x, y\big)\,; \tag{4.3}$$

hence the parametrization of the geodesics ending at 0 is given by

$$\eta^b(t) \;=\; T \circ \gamma_v^b(t) \;=\; \frac{1}{b^2+e^{2t}}\big(-b, e^t\big)\,.$$

We obtain now the parametrization of the geodesics ending at any point $a \in \mathbb{R}$ by composing $\eta^b(t)$ with the translation $T_a(z) = z + a$, more precisely

$$\eta_a^b(t) \;=\; T_a \circ \eta^b(t) \;=\; \left(a - \frac{b}{b^2+e^{2t}}, \frac{e^t}{b^2+e^{2t}}\right). \tag{4.4}$$

4.3 The Horocycle flows

In the previous section we spoke about the geodesic flow on the hyperbolic plane and on its quotients; here we want to introduce two other important flows on $\mathbb{H}$ (resp. on its quotients) that will be, together with the geodesic flow, the central point of chapter 5. Observe that any hyperbolic circle $C(z,r)$ is an Euclidean circle in $\mathbb{H}$ (although its center is not z) as can be seen applying the standard isometry between $\mathbb{H}$ and Δ (see the appendix). Thus if $v \in T_z\mathbb{H}$ is a unit tangent vector the circle $C(\gamma_v(r), r)$, to which $\gamma_v(0) = z$ always belongs (because the geodesic is parametrized by arc-length), converges as r increases to an Euclidean circle touching $\partial\mathbb{H}$ in γ_{v+} and having v as inward normal at z. This circle is called the *positive horosphere* $S^+(v)$; observe that in the case $\gamma_{v+} = \infty$ the positive horosphere is the horizontal line through z. Similarly the circle $C(\gamma_v(-r), r)$ converges as r increases to the Euclidean circle $S^-(v)$, called the *negative horosphere*, touching $\partial\mathbb{H}$ at γ_{v-} and having v as outward normal at z (see figure 4.2). Notice that horospheres are not geodesics and viceversa, since horospheres are tangent to $\partial\mathbb{H}$ while geodesics are orthogonal to the boundary; moreover all the inward-pointing normals to $S^+(v)$ (resp. all the outward-pointing normals to $S^-(v)$) define geodesics with the same point at ∞ $(-\infty)$, that is γ_{v+} (γ_{v-}). The *horocycle flow* on $T_1\mathbb{H}$ is the flow

$$h_t^* : T_1\mathbb{H} \longrightarrow T_1\mathbb{H}$$

which slides counterclockwise the inward normal vectors to each $S^+(v)$ along $S^+(v)$ at unit speed; if $\pi : T_1\mathbb{H} \longrightarrow \mathbb{H}$ denotes the canonical projection then

$$\frac{d}{dt}\pi h_t^*(v) \;\perp\; h_t^*(v)$$

is a unit vector and the distance from $\pi(v)$ and $\pi(h_t^*(v))$ measured along $S^+(v)$ is t. The curve

$$\{h_t^*(v) \mid t \in \mathbb{R}\}$$

of inward-pointing vectors normal to $S^+(v)$ is the *horocycle* in $T_1\mathbb{H}$, as distinct from the curve $S^+(v)$ in $\mathbb{H}$ which is the horosphere; observe that, since any inward pointing vector to $S^+(v)$ defines a geodesic with the same point at infinity, we have $\gamma_{h_t^*(v)+} = \gamma_{v+}$ for any $t \in \mathbb{R}$. There is another horocycle flow $h_t : T_1\mathbb{H} \longrightarrow T_1\mathbb{H}$ which slides clockwise normal outward pointing vectors to $S^-(v)$ along $S^-(v)$ at unit speed; in this case $\gamma_{h_t(v)-} = \gamma_{v-}$ for any $t \in \mathbb{R}$.

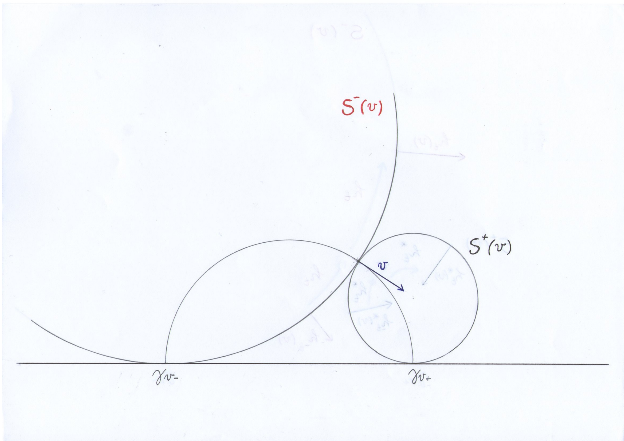


Figure 4.1: The horospheres $S^+(v)$ and $S^-(v)$ associated to a vector $v \in T_1\mathbb{H}$.

Remark. *Since $S^+(v) = S^-(-v)$ and sliding a vector $v \in T_1\mathbb{H}$ to the right along $S^+(v)$ is equivalent to slide $-v$ to the left along $S^-(-v)$ and then reverse the obtained vector, we get the relation*

$$h_t^*(v) = -h_{-t}(-v)\,. \tag{4.5}$$

Observe that the geodesic flow g_t maps the horizontal line through i (that is the horosphere $S^+(i_i)$) into the horizontal line through ie^t (that is the horosphere $S^+(ie^t_{ie^t})$); hence the hyperbolic length along the horocycle h_t^* decreases by a factor e^{-t} while it increases by a factor e^t along the horocycle h_t. In other words we have the following relations between g_t, h_t^* and h_t

$$\begin{cases} g_t \circ h_s^* &= h_{se^{-t}}^* \circ g_t\,; \\ g_t \circ h_s &= h_{se^t} \circ g_t\,; \end{cases} \tag{4.6}$$

At each point $v \in T_1\mathbb{H}$ we have a three-dimensional tangent space $T_vT_1\mathbb{H}$ with a basis given by the vectors tangent, at $t = 0$, to the curves $g_t(v), h_t^*(v)$ and $h_t(v)$ in $T_1\mathbb{H}$. The one-dimensional subspaces of $T_vT_1\mathbb{H}$ spanned by these vectors will be called E_v^0, E_v^s and E_v^u; they are invariant under the geodesic flow in the sense that $T_{g_t}(E_v^s) = E_{g_t(v)}^s$, ... Now the partitions of $T_1\mathbb{H}$ into g_t, h_t^* and h_t-orbits are invariant under left-multiplication and so is the splitting

$$T_vT_1\mathbb{H} \;=\; E_v^s \oplus E_v^0 \oplus E_v^u\,; \tag{4.7}$$

Since any inner product on $T_{i_i}T_1\mathbb{H}$ extends to a left-invariant riemannian metric on $T_1\mathbb{H}$, combining relations in (4.6) with $g_t \circ g_s = g_s \circ g_t$ we get

$$\|T_{g_t}|E_v^0\| \;=\; 1\,, \quad \|T_{g_t}|E_v^s\| \;=\; e^{-t}\,, \quad \|T_{g_t}|E_v^u\| \;=\; e^t\,; \tag{4.8}$$

Any flow with this property, that is the tangent bundle splits as direct sum of three invariant subbundles, one tangent to the orbits, one that is contracted at some exponential rate in t and one that is expanded at some exponential rate in t, is called *hyperbolic flow* or *Anosov flow*. In our case E_v^s, E_v^u are the tangent spaces to the positive and negative horocycles, which we shall call $W^{ss}(v)$

and $W^{uu}(v)$, the *strong stable* and *strong unstable* manifold of v in $T_1\mathbb{H}$. These manifolds form two partitions of $T_1\mathbb{H}$ with the property

$$v' \in W^{ss}(v) \quad \Longleftrightarrow \quad d(g_t(v'), g_t(v)) \longrightarrow 0 \ \text{ as } \ t \longrightarrow +\infty$$

$$v' \in W^{uu}(v) \quad \Longleftrightarrow \quad d(g_t(v'), g_t(v)) \longrightarrow 0 \ \text{ as } \ t \longrightarrow -\infty$$

Now we want to see how the horocycle flows can be written using the identification $T_1\mathbb{H} \cong PSL(2,\mathbb{R})$; since $S^+(i_i)$ is the horizontal line through i we have that $h_t^*(i_i) = i_{i+t}$ for all $t \in \mathbb{R}$. Therefore in $PSL(2,\mathbb{R})$ the horocycle flow is given by

$$h_t^*(I) \ = \ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} =: H_t^*$$

and more generally $h_t^*(B) = BH_t^*$ for all $B \in PSL(2,\mathbb{R})$. Using (4.5) and the fact that reversing a vector corresponds in $PSL(2,\mathbb{R})$ to replacing B with BJ (J is the 2×2-symplectic identity) we get

$$h_t : PSL(2,\mathbb{R}) \longrightarrow PSL(2,\mathbb{R}) \, , \quad h_t(B) = BJH_t^*J = B \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} = BH_t \, .$$

The horospheres are preserved by the action of $PSL(2,\mathbb{R})$ on $\mathbb{H}$, indeed any circle is mapped by an element of $PSL(2,\mathbb{R})$ into a circle and the tangency condition is also preserved; so, quotienting out by the action of a group $\Gamma < PSL(2,\mathbb{R})$ of parabolic and hyperbolic isometries acting properly discontinuously on $\mathbb{H}$, gives us the positive and negative horosphere $S^+(v)$, $S^-(v)$ for a unit vector $v \in T_1M$, where $M = \Gamma\backslash\mathbb{H}$ is a complete hyperbolic surface. Moreover since horocycles are preserved by the isometries in Γ the horocycle flows project to h_t^*, $h_t : T_1M \longrightarrow T_1M$ (see figure 4.3). The unit tangent bundle T_1M can be seen also as $\Gamma\backslash PSL(2,\mathbb{R})$ and with this identification the horocycle flows on T_1M are given by

$$h_t^*(\Gamma B) = \Gamma B H_t^* \, , \qquad h_t(\Gamma B) = \Gamma B H_t \, .$$

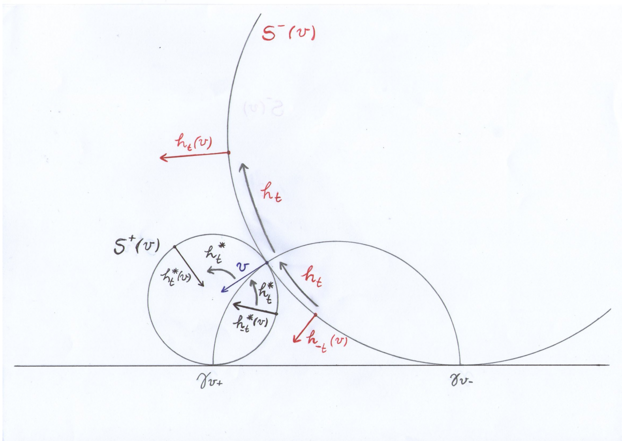


Figure 4.2: The horocycle flows in $T^1\mathbb{H}$; since horocycles are preserved by the isometries in Γ these horocycle flows project to $h_t^*, h_t : T_1M \longrightarrow T_1M$.

We end this section computing the (clockwise) arclength parametrisation of the horospheres; as done already for the geodesics, we start from the parametrisation of the horospheres tangent to the boundary at ∞ obtaining then the parametrisation of horospheres tangent to any point of the boundary by applying a suitable isometry of $\mathbb{H}$. One can easily check that the parametrisation of the horizontal lines (i.e. the horospheres tangent at ∞) is given by

$$\gamma_b(t) = (-b\,t, b) \, ;$$

therefore the parametrisation of the horospheres tangent to $\partial\mathbb{H}$ at the origin is obtained by composing γ_b with the isometry T in (4.3), more precisely

$$\eta_b(t) \;=\; T \circ \gamma_b(t) \;=\; \frac{1}{b(1+t^2)}\left(t, 1\right).$$

Finally, the horospheres tangent to $a \in \mathbb{R}$ are parametrised by

$$\eta_b^a(t) \;=\; \left(a + \frac{t}{b(1+t^2)}, \frac{1}{b(1+t^2)}\right) \tag{4.9}$$

as we get applying the translation $T^a(z) = z + a$ to η_b.

4.4 Recurrence and denseness via symbolic dynamics

We can study the geodesic and horocycles flows either in $T_1\mathbb{H}$ or in T_1M (which has finite measure); although in $T_1\mathbb{H}$ it is easier to understand the various orbits and their attracting and repelling properties, it is in T_1M that the interesting recurrent dynamic and close (that is periodic) or even dense orbits occur. We only enounce these properties in $T_1\mathbb{H}$ and try to explain analogous behaviours for other (easier) flows, which take place in Euclidean rather than Hyperbolic spaces. However at the end of the section, following the original from Hedlund contained in [21], we will give shortly a (unfortunately non completely detailed) proof of theorem 4.4.2.

Theorem 4.4.1. *If $M = \Gamma \setminus \mathbb{H}$ is a compact manifold then the periodic orbits of the geodesic flow g_t are dense in T_1M; moreover g_t has a orbit which is dense in T_1M.*

Theorem 4.4.2 (Hedlund)**.** *If $M = \Gamma \setminus \mathbb{H}$ is compact then each orbit of h_t^* and h_t is dense in T_1M.*

We refer to [20], or to the original from Hedlund, for a complete and rigorous proof of these results; observe that theorem 4.4.2 does not hold if M is finite-area but not compact. Indeed in this case the discrete group of isometries $\Gamma < PSL(2,\mathbb{R})$ (defining M) contains a parabolic element that, without loss of generality (see [20]), is of the form $z \longmapsto z + a$ for some $a \neq 0$. Since a horocycle at ∞ corresponds to a horizontal line at some height y we get that such a horocycle is periodic for h_t^* with period a/y and hence it is not dense. Now we want to describe analogous properties as those in theorem 4.4.1 and 4.4.2 in an easier case in order to understand and to visualize better such contrasting and somehow unexpected behaviours; we will consider two flows, the first one ψ_t will be the suspension of a hyperbolic automorphism of $\mathbb{T}^2$ into itself while the second one φ_t will be an irrational translation flow on $\mathbb{T}^2$. Each of those flows can be studied on the universal cover $\mathbb{R}^2$, where orbits are easier to see but where there is no recurrence.

Example 4.1. *If β is a real number define the flow*

$$\varphi_t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2\,, \qquad (x,y) \longmapsto (x+t, y+\beta t)\,;$$

all orbits are parallel lines with slope β. Moreover since φ_t commutes with the action of $\mathbb{Z}^2$ on $\mathbb{R}^2$ it induces a flow on $\mathbb{T}^2$ that we also denote by φ_t; observe that if β is a rational number then each orbit in $\mathbb{T}^2$ is periodic, indeed if we write $\beta = \frac{p}{q}$ with p, q relatively prime and $q > 0$ then

$$\varphi_q(x,y) = (x+q, y+p) = (x,y)$$

in $\mathbb{T}^2$. Thus in this case the torus is covered by closed orbits; however this is not the case we are interested in, indeed is in the irrational case that a behaviour like that of the horocycle flows takes place. More precisely if β is irrational then each orbit of the flow φ_t in $\mathbb{T}^2$ is non closed and dense (this fact was observed for the first time by Jacobi in 1835), indeed it meets the circle $\{(x,y) \in \mathbb{T}^2 \mid x = 0\}$ in the dense subset

$$\{(0, a + m\beta\ (mod\ 1)) \mid m \in \mathbb{Z}\}$$

for a suitable $a \in (0,1)$.

Example 4.2. *Define $\psi_t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ to be the translation along the third coordinate, that is $(x, y, z) \longmapsto (x, y, z + t)$; this flow on $\mathbb{R}^3$ commutes with f_1, f_2 and f_3 given by*

$$(x, y, z) \xmapsto{f_1} (x + 1, y, z)\,, \quad (x, y, z) \xmapsto{f_2} (x, y + 1, z)\,, \quad (x, y, z) \xmapsto{f_3} (x + y, x, z - 1)$$

and hence with the group generated. Observe that $\mathbb{R}^2$ quotiented by the action of $< f_1, f_2 >$ (of course f_1 and f_2 are thought restricted to the first two variables) is the 2-torus $\mathbb{T}^2$; now since ψ_t commutes with $< f_1, f_2, f_3 >$ it induces a flow on the quotient N of $\mathbb{R}^3$

$$N \;=\; \mathbb{T}^2 \times [0,1] \Big/ \big[(x, y, 1) \sim (x + y, x, 0)\big]\,.$$

The flow $\psi_t : N \longrightarrow N$ is called the suspension of the map $A : \mathbb{T}^2 \longrightarrow \mathbb{T}^2$ defined by the matrix

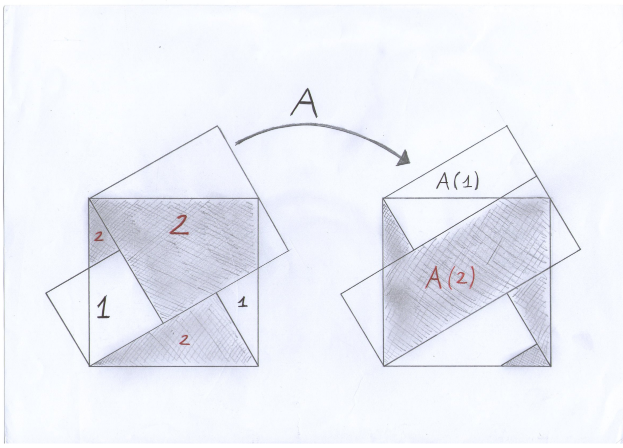
$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\,, \qquad A : (x, y) \longmapsto (x + y, x)\,;$$

it is the flow on the direction of the second factor of N. Orbits of ψ_t can be understood by their successive intersections with $\mathbb{T}^2 \times \{0\}$ (these points form an orbit of A); an orbit $\{A^n(p) \mid n \in \mathbb{Z}\}$ can be seen using symbolic dynamics (see [20] for further details) which records the sequence $k = \{k_n\}_{n\in\mathbb{Z}} \subseteq \{1, 2\}^{\mathbb{Z}}$ of rectangles 1 or 2 (see the figure below) to which the sequence $A^n(p)$ belongs. Thus for all $n \in \mathbb{Z}$ we have $k_n = 1$ if and only if $A^n(p) \in 1$ and similarly $k_n = 2$ if and only if $A^n(p) \in 2$; the edges of 1 and 2 are intervals I^u, I^s which lie in the expanding and contracting eigenspace E^u, E^v (corresponding to the eigenvalues of the matrix associated to A) through the fixed point $(0, 0)$; therefore we get

$$A(I^u) \supseteq I^u\,, \qquad A(I^s) \subseteq I^s\,.$$

In fact $A(1)$ crosses 2 while $A(2)$ crosses both, hence in k we can not have two successive 1 but any other combination is possible. Define the set $\Sigma_A := \;\{k \mid k_n \in \{1, 2\}\,,\; k_n k_{n+1} \neq 11 \;\forall\, n \in \mathbb{Z}\}$ and

$$\sigma_A : \Sigma_A \longrightarrow \Sigma_A, \quad (\sigma k)_n = k_{n+1}\,, \qquad \pi : \Sigma_A \longrightarrow \mathbb{T}^2, \quad \pi(k) = \bigcap_{n\in\mathbb{Z}} A^{-n}(k_n)\,.$$



By definition we have $\pi \cdot \sigma_A = A \cdot \pi$; moreover π takes the σ_A-orbit of each sequence k to an A-orbit in $\mathbb{T}^2$. Here we have ignored the ambiguity over which symbol sequence attach to a point p whose orbit meets $I^s \cap I^u$ (see [17] for further details); however, a part from points in $I^u \cap I^s$ which can be represented by two or four symbol sequences there is a 1-1 correspondence between symbol sequences and points in $\mathbb{T}^2$. This general technique, introduced by Adler and Weiss in [18] and by Sinai in [19], is called the method of Markov partitions. Cross-sections to the flow are partitioned into rectangles which are product of pieces of stable and unstable manifolds chosen so that the first return map to the cross sections maps each rectangle exactly across several rectangles in the unstable direction. A symbol space like Σ_A and a time of first return function (in this case the constant function 1) suffice to describe the flow orbits. The flow ψ_t has two contrasting properties that are easily established using the symbolic dynamics:

- *Its periodic orbits are dense in N.*
- *It has a dense orbit in N.*

Indeed it suffices to show that periodic orbits of A are dense in $\mathbb{T}^2$ and that A has a dense orbit in $\mathbb{T}^2$; it is easy to construct periodic sequences in Σ_A which repeat $k_{-m} \cdot ... \cdot k_0 \cdot ... \cdot k_m$ indefinitely and this corresponds to periodic orbits of A which are close, for large m, to any given $\pi(k)$ in $\mathbb{T}^2$. To have an orbit dense in Σ_A corresponding to an A-orbit dense in $\mathbb{T}^2$ a sequence k must contain somewhere all blocks of symbols 1 and 2 of length r (without two successive 1's!!!) for each r; it is easy to write down such a sequence be interposing a 2 where necessary between two successive 1's.

If $M = \Gamma\backslash\mathbb{H}$ is a compact manifold a similar, though more complicated, construction can be made for the geodesic flow $g_t : T_1M \longrightarrow T_1M$ and this proves theorem 4.4.1. Thus $\psi_t : N \longrightarrow N$, which likes the geodesic flow g_t, has its periodic orbits dense, a dense orbit and is an Anosov flow with directions of exponential repultion and attraction transverse to each orbit. By contrast $\varphi_t : \mathbb{T}^2 \longrightarrow \mathbb{T}^2$, which likes the horocycle flows h_t^* and h_t, has every orbit dense and so no periodic orbits. Now we want briefly explain how originally Hedlund in 1936 proved theorem 4.4.2; the proof consists in several tricky (but quite easy to prove) steps. Rather than bore the reader describing all of them we have decided to focus on the most important ones and refer to [20] for the rest. Hereafter we suppose that M is a compact manifold obtained by quotenting out the unit disc Δ by a group of discrete isometries (this is clearly equivalent to see M as a quotient of the hyperbolic plane $\mathbb{H}$); hence let $M = \Gamma \setminus \Delta$ be compact and let $q \in \Delta$ be a point. The first step of the proof is given by the following

Lemma 4.4.3. *Let $I \subseteq \partial\Delta$ be an interval such that for any $J \subseteq I$ the set*

$$\Gamma \left[\bigcup \left\{ h_{\mathbb{R}^+}^*(v) \mid v \in T_q\Delta\,,\ \gamma_{v+} \in J \right\} \right] \tag{4.10}$$

is dense in T_1M; then there exists an infinite subset K which is dense in I and such that $\Gamma\,[h_{\mathbb{R}^+}^*(v)]$ is dense in T_1M for each $v \in T_q\Delta$ with $\gamma_{v+} \in K$.

Proof. Let $\{U_n\}_{n\in\mathbb{N}}$ be a sequence of subset of $T_1\Delta$ such that $\{\Gamma U_n\}_{n\in\mathbb{N}}$ forms a basis for the topology on T_1M. Suppose $J \subseteq I$ fixed, then (4.10) implies that there exists $v_1 \in T_q\Delta$ such that

$$\Gamma h_{\mathbb{R}^+}^*(v_1) \ \cap\ \Gamma U_1 \ \neq\ \emptyset\,, \qquad \gamma_{v_1+} \in J\,;$$

Let V_1 be a small interval in J containing γ_{v_1+} so that the above intersection is still non-empty for any $v \in T_q\Delta$ with $\gamma_{v+} \in V_1$; since (4.10) holds for any interval contained in I we can find a vector $v_2 \in T_q\Delta$ such that $\Gamma\, h_{\mathbb{R}^+}^*(v_2)$ intersects both ΓU_1 and ΓU_2. Repeating the same process we get $v \in T_q\Delta$ such that $\gamma_{v+} \in \cap_n V_n$ and $\Gamma\, h_{\mathbb{R}^+}^*(v)$ meets each ΓU_n in T_1M and so is dense there; since such points γ_{v+} can be found in any J they form an infinite dense subset of I. $\square$

It is an easy consequence of theorem 4.4.1 (in particular of the fact that g_t-periodic orbits are dense in T_1M) that for any fixed interval $J \subseteq \partial\Delta$ the set in (4.10) is dense in T_1M; the reader can find the proof and more details about that in [20]. Furthermore it follows immediately from

$$g_t \cdot h_s^* \ =\ h_{se^{-t}}^* \cdot g_t$$

that if a point $k \in \partial\Delta$ has $\Gamma h_{\mathbb{R}^+}^*(v)$ dense in T_1M for some v approaching k (i.e. such that $\gamma_{v+} = k$) then the same is true for any other $w \in T_1M$ apporaching k. Thus so far we have proved that at least for a dense set of points K in $\partial\Delta$ all horocycles approaching K are dense in T_1M; therefore theorem 4.4.2 is proved for a dense subset in $\partial\Delta$. Now it remains to show the same for any other point that is not in this dense subset; as preliminary result one can prove that the thesis is true if $k \in \partial\Delta$ is a fixed point of some hyperbolic isometry $B \in \Gamma$ and with this fact get the last step.

Lemma 4.4.4. *If C is a horocycle in Δ with copies AC (for various $A \in \Gamma$) of Euclidean radius arbitrarily close to 1 the ΓC is dense in T_1M.*

Proof. Take $v \in T_1\Delta$ with $\Gamma v \in T_1M$ fixed by g_ω for some $\omega > 0$ and let $B \in \Gamma$ s.t. $v \longmapsto g_w(v)$; denote by $k \in \partial\Delta$ the point approached by the horocycle C and take a sequence $\{A_nC\}$ of copies of C whose Euclidean radius increases to 1. The horospheres $\pi(A_n C)$ in Δ cut $\gamma_v(\mathbb{R}^+)$ at one point with an angle which approaches $\frac{\pi}{2}$. Now pick m_n such that

$$B^{m_n}\big(\gamma_v(\mathbb{R}^+) \cap A_n C\big) \;\in\; \gamma_v([0,w])\,;$$

these points subconverge (i.e. converge up to take a subsequence) to some point of $\gamma_v([0,w])$ where the angle of intersection approaches $\frac{\pi}{2}$, so $B^{m_n} A_n k \longrightarrow \gamma_{v-}$. The horocycles $B^{m_n} A_n C$ subconverge to a horocycle touching $\partial\Delta$ at γ_{v-} and this is dense because γ_{v-} is left fixed by B; since all of these horocycles are the same in $T_1 M$ we get that ΓC is dense. □

Now we are finally ready to complete the proof of theorem 4.4.2; let us consider a fundamental domain (see appendix) $V \subseteq \Delta$ which contains the origin and put

$$y := \;\sup\,\{d(0,z) \mid z \in V\}\,.$$

Let $C = h^*_{\mathbb{R}}(v)$ be a horocycle with $v \in T_0\Delta$ unit vector; observe that

$$d(g_t(v), h^*_s(v)) \;\geq\; t \tag{4.11}$$

holds for every $s \in \mathbb{R}$. This fact is easier to see in the hyperbolic plane $\mathbb{H}$; it suffices to prove the inequality in the case $v = i_i$ and then display this to any other vector $w \in T_1\mathbb{H}$ using a suitable isometry of $\mathbb{H}$. If $v = i_i$ the geodesic γ_{i_i} is the imaginary axis parametrized by arc-length while the horosphere is the horizontal line at height i and hence the inequality is clear; therefore (4.10) holds for any unit tangen vector $v \in T_1\mathbb{H}$ and since $\mathbb{H}$ and Δ are isometrically equivalent we get (4.10) also for the unit disc Δ. If we choose now $B_t \in \Gamma$ such that $B_t(\pi g_t(v)) \in V$ then

$$B_t(\pi C) \;\subseteq\; \{z \in \Delta \mid d(0,z) \geq t - y\}\,;$$

this exhibits copies of C of Euclidean radius as near to 1 as we please by choosing large values of t so that ΓC is dense in $T_1 M$ be lemma 4.4.4 and this completes the proof.

Chapter 5

Horocycle flows from the Hamiltonian viewpoint

5.1 Horocycle flows for the hyperbolic plane $\mathbb{H}$

Let $\mathbb{H}$ be the hyperbolic plane and let $T\mathbb{H}$, $T^*\mathbb{H}$ be its tangent, resp. cotangent, bundle; denote by $z = (x, y)$ the standard coodinates on $\mathbb{H}$ and if $v \in T_z\mathbb{H}$, $p \in T_z^*\mathbb{H}$, is a tangent, resp. cotangent, vector then denote by $v = (v_x, v_y)$, resp. $p = (p_x, p_y)$, its components. Define the Lagrangian

$$L(x, v) \;=\; \frac{1}{2}\|v\|_x^2 + \eta_x(v) \tag{5.1}$$

where $\|v\|_x := \sqrt{g_x(v, v)}$ is the norm induced by g and η is the 1-form on $\mathbb{H}$ given by

$$\eta_{(x,y)} := \frac{dx}{y}$$

for all $(x, y) \in \mathbb{H}$. The Lagrangian defined in (5.1) satisfies the properties of definition 1.1.1 because the 1-form η is bounded on $\mathbb{H}$ with respect to the metric g; furthermore the differential of η is the area form σ on $\mathbb{H}$ induced by the metric g, indeed

$$d\eta \;=\; -\frac{1}{y^2}\, dy \wedge dx \;=\; \frac{1}{y^2}\, dx \wedge dy\,.$$

Since the area form σ is exact on $\mathbb{H}$ we are allowed to study the horocycle flow using the Lagrangian formalism; this will be no longer true for the quotients of the hyperbolic plane (that is for the hyperbolic surfaces). Therefore in those cases we shall use the Hamiltonian viewpoint, since we can not write the Lagrangian any more. The Lagrangian in (5.1) is just a particular case of magnetic Lagrangian with no potential energy; thus its energy function is

$$E(z, v) \;=\; \frac{1}{2}\,\|v\|_z^2 \tag{5.2}$$

and the associated Euler-Lagrange equation is given by

$$\frac{D}{dt}\,\dot{z} \;=\; Y_z(\dot{z})$$

where $\frac{D}{dt}$ denotes the covariant derivative and $Y : T\mathbb{H} \longrightarrow T\mathbb{H}$ is the bundle map defined by

$$d\eta_z(u, v) \;=\; g_z(Y_z(u), v) \tag{5.3}$$

for all $z \in \mathbb{H}$ and for all $u, v \in T_x\mathbb{H}$. If we denote by (u_x, u_y), (v_x, v_y), (w_x, w_y) respectively the components of u, v, $Y_z(u)$ then by (5.3) we get that

$$\frac{1}{y^2}\,(u_x v_y - u_y v_x) \;=\; \frac{1}{y^2}\,(w_x v_x + w_z v_z)$$

for any possible choiche of u_x, u_y, v_x, v_y; therefore applying this equality for $v_x = 1$ and $v_y = 0$ (resp. for $v_x = 0$ and $v_y = 1$) we get that $w_x = -u_y$ (resp. $w_y = u_x$), so that $Y_x(u) = J \cdot u$, where J is the 2×2-simplectic identity. Hence

$$\frac{D}{dt}\,\dot z \;=\; J \cdot \dot z$$

is the Euler-Lagrangian equation associated to the Lagrangian L in (5.1). Since $\mathbb{H}$ is an open subset of $\mathbb{R}^2$ the Euler-Lagrange equation can also be written in coordinates; if $((x,y),(v_x,v_y))$ is the standard coordinate system in $T\mathbb{H}$ then the Lagrangian L can be written as

$$L\big((x,y),(v_x,v_y)\big) \;=\; \frac{1}{2y^2}\left(v_x^2+v_y^2\right) + \frac{v_x}{y}\,. \tag{5.4}$$

Thus the partial derivatives of L are

$$\begin{cases} \dfrac{\partial L}{\partial x} \;=\; 0\,; \\[2ex] \dfrac{\partial L}{\partial y} \;=\; -\dfrac{1}{y^3}\big(v_x^2+v_y^2\big) - \dfrac{v_x}{y^2}\,; \end{cases} \qquad\qquad \begin{cases} \dfrac{\partial L}{\partial v_x} \;=\; \dfrac{v_x}{y^2} + \dfrac{1}{y}\,; \\[2ex] \dfrac{\partial L}{\partial v_y} \;=\; \dfrac{v_y}{y}\,; \end{cases}$$

and hence the Euler-Lagrange equation is given by

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{v_x}{y^2} + \dfrac{1}{y}\right) = 0\,; \\[2ex] \dfrac{d}{dt}\left(\dfrac{v_y}{y^2}\right) = -\dfrac{1}{y^3}\big(v_x^2+v_y^2\big) - \dfrac{v_x}{y^2}\,; \end{cases} \tag{5.5}$$

It is clear from the system above that the momentum associated to x is a prime integral, that is it is constant along the motion; one could also deduce this fact observing that the Lagrangian L does not depend explictly on x. Since the Lagrangian is autonomous, i.e. not time dependent, we get that also the energy function is a prime integral; so we have two prime integrals. We will use these prime integrals in the following section to study the dynamics for low energy levels.

Theorem 5.1.1. *Let $L : T\mathbb{H} \longrightarrow \mathbb{R}$ be the Lagrangian in (5.4), then the Mañé's critical value of L is $c(L) = \frac{1}{2}$ and it coincides with all the other critical values introduced in chapter 2. Moreover the Aubry set is empty, hence there are no minimizing measures, and the Euler-Lagrange flow at the energy level $E \equiv \frac{1}{2}$ is the horocycle flow.*

Proof. By (5.9) we get immediately that the curves $\dot x = -y$, $\dot y = 0$ are solutions of the Euler-Lagrange equation. The images of these curves are the horizontal lines parametrized by arc-length, i.e. the stable horospheres associated to vertical geodesics of the hyperbolic plane. The energy of such solutions is $\frac{1}{2}$, indeed

$$E\big((x,y),(-y,0)\big) \;=\; \frac{1}{2}\cdot\|(-y,0)\|_z^2 \;=\; \frac{1}{2}\,.$$

We prove now that the energy level $\frac{1}{2}$ is critical (in the sense of Mañé); therefore there is a drastic change in the dynamic crossing the energy value $\frac{1}{2}$. In particular orbits with energy lower than $\frac{1}{2}$ will be closed (see section 5.2) while the flow at the energy level $E = \frac{1}{2}$ is the horocycle flow, hence there are no closed orbits. Finally, the Euler-Lagrange flow for supercritical energy level $k > \frac{1}{2}$ is the reparametrization of the geodesic flow for a suitable Finsler metric. First we show that $c(L) \le \frac{1}{2}$; if $v = (v_x, v_y)$ is a tangent vector we have

$$L(z,v) \;=\; \frac{1}{2}\,\|v\|_z^2 + \eta_z(v) \;\ge\; \frac{1}{2}\,\|v\|_z^2 - |\eta_z(v)| \;=\; \frac{1}{2}\,\|v\|_z^2 - \frac{1}{y}\,|v_x| \;\ge\; \frac{1}{2}\,\|v\|_z^2 - \|v\|_z \;\ge -\frac{1}{2}\,.$$

Therefore $L+\frac{1}{2} \ge 0$ and this implies $c(L) \le \frac{1}{2}$; conversely consider a curve γ_r parametrizing clockwise the boundary of a geodesic disc D_r of radius $r > 0$ with constant speed $\|\dot\gamma_r(t)\|_{\gamma_r(t)} \equiv a > 0$. Since the energy of γ_r is constant and equal to $E(\gamma_r) = \frac{1}{2}\,a^2$, we get

$$
\begin{aligned}
\int_{\gamma_r}\left(L+\frac{1}{2}a^2\right)d\lambda \;&=\; \int_{\gamma_r} L_v\cdot v\,d\lambda \;=\int_{\gamma_r}\left(v+\frac{1}{y}\,e_x\right)\cdot v\;d\lambda \;=\\
&=\; \int_{\gamma_r}\left(\|v\|_z^2+\eta_z(v)\right)d\lambda \;=\int_{\gamma_r}\|v\|_z^2\,d\lambda \;-\int_{D_r}\sigma \;=\\
&=\; a^2\cdot\text{length}\,(\gamma_r) \;-\; \text{area}\,(D_r) \;=\; 2\pi a^2\sinh(r)-4\pi\sinh^2\!\left(\frac{r}{2}\right) \;=\\
&=\; 2\pi\left[\frac{1}{2}(a^2-1)e^r-e^{-r}\right]+2\pi
\end{aligned}
$$

where the fourth equality follows from Stokes theorem and the penultimate from the formulas (B.8) and (B.9) in the appendix. From the calculation above we get that if $a<1$ then for $r>0$ big enough the first integral is negative; therefore if $a<1$ we can always find a closed (and absolutely continuous) curve with negative $(L+\frac{1}{2}a^2)$-action and this implies that $c(L)\geq\frac{1}{2}$. Thus the Mañé critical value is $c(L)=\frac{1}{2}$; observe that since $\mathbb{H}$ is simply connected all the critical values introduced in chapter 2 are equal, that is

$$
c_u(L) \;=\; c_a(L) \;=\; c_0(L) \;=\; c(L) \;=\; \frac{1}{2}\,.
$$

Remark. *Consider two points $x,y\in\mathbb{H}$ such that $x_2=y_2$ and consider the curve γ parametrizing by arc-length the horizontal segment joining x to y, i.e. such that in any point of the segment the tangent vector is $(-x_2,0)$. Observe that $L(\gamma,\dot\gamma)+\frac{1}{2}=0$ for any point of the segment, so that*

$$
\mathbb{A}_{L+\frac{1}{2}}(\gamma) \;=\int\left[L(\gamma,\dot\gamma)+\frac{1}{2}\right] \;=\; 0\,.
$$

On the other hand any other curve η joining x to y must have non negative $(L+\frac{1}{2})$-action (because $L+\frac{1}{2}\geq 0$); therefore γ is a semistatic curve and in general the whole horizontal horosphere parametrized by arclength is semistatic, that is any vector $((x_1,x_2),(-x_2,0))\in\Sigma(L)$.

We show now that the solutions of the Euler-Lagrange equation at energy level $\frac{1}{2}$ are the horospheres parametrised by arclength and that the Aubry set $\widehat{\Sigma}(L)$ defined in section 2.3 is empty; therefore there are no minimizing measures. Observe that this is not in contradiction with theorem 2.4.2 since $\mathbb{H}$ not compact; hence this is an example of Lagrangian that does not admit minimizing measures and for which the Aubry set is empty. So let $T:\mathbb{H}\longrightarrow\mathbb{H}$ be an isometry; since

$$
d(T^*\eta) \;=\; T^*(d\eta) \;=\; T^*\sigma \;=\; \sigma \;=\; d\eta
$$

we get that the form $T^*\eta-\eta$ is closed and hence exact, that is there exists a smooth function f such that $df=T^*\eta-\eta$. Thus if $t,w\in\mathbb{H}$ are two points, for any curve $\gamma\in C(t,w)$ we have

$$
\begin{aligned}
\mathbb{A}_L(T\circ\gamma) \;&=\; \int L(T\circ\gamma,dT(\gamma)[\dot\gamma]) \;=\int\left[\frac{1}{2}\left\|dT(\gamma)[\dot\gamma]\right\|^2_{T(z)}+\eta_{(T\circ\gamma)}(dT(\gamma)[\dot\gamma])\right] \;=\\
&=\; \int\left[\frac{1}{2}\|\dot\gamma\|_z^2+(T^*\eta)_\gamma(\dot\gamma)\right] \;=\int\left[\frac{1}{2}\|\dot\gamma\|_z^2+\eta_\gamma(\dot\gamma)+df(\gamma)[\dot\gamma]\right] \;=\\
&=\; \int\left[\frac{1}{2}\|\dot\gamma\|_z^2+\eta_\gamma(\dot\gamma)\right]+f(w)-f(t) \;=\; \mathbb{A}_L(\gamma)+f(w)-f(t)\,.
\end{aligned}
$$

Hence for any isometry T of $\mathbb{H}$ and for any pair t,w of points of $\mathbb{H}$, the difference $\mathbb{A}_L-\mathbb{A}_L\circ T$ is constant and depends only on T and on t,w; in particular we get that the property of a curve to be minimizer of the action potential is preserved by the isometries of the hyperbolic plane. In other words $\Sigma(L)$ and $\widehat{\Sigma}(L)$ are invariant under dT; therefore, since the horizontal horospheres are solutions for the Euler-Lagrange equation and the isometries of $\mathbb{H}$ are transitive over $T_1\mathbb{H}$ (in the sense explained in chapter 4) we get that the horospheres parametrized by arc-length are solutions of the Euler-Lagrange equation. On the other hand any solution of the Euler-Lagrange equation with energy $\frac{1}{2}$ must be a horosphere parametrized by arc-length, indeed for any point $(z,v)\in T_1\mathbb{H}$ there is a (unique) horosphere through (z, v), which is solution of the Euler-Lagrange equation; hence non horosphere solution would contradict the uniqueness of the solution. Furthermore horospheres can not be static because any horosphere h_1 can be sent by an isometry to another horosphere h_2 such that $h_1 \cap h_2 \neq \emptyset$ and this would contradict the graph property in theorem 2.3.5; hence $\widehat{\Sigma}(L) = \emptyset$. □

Given the Lagrangian L in (5.1) it is easy to write down the corresponding Hamiltonian using the Legendre transform

$$H(z, L_v) \;=\; H \circ \mathcal{L}(z, v) \;= < L_v, v >_z - L(z, v) \;=\; E(x, v)\,;$$

since $L_v = v + \eta$ and $E(z, v) = \dfrac{1}{2}\,\|v\|_z^2$, we get that

$$H(z, p) \;=\; \frac{1}{2}\|p - \eta\|_z^2 \tag{5.6}$$

where for sake of simplicity we denote by $\|\cdot\|_z$ both the norm on $T\mathbb{H}$ and the dual norm on $T^*\mathbb{H}$. If $((x, y), (p_x, p_y))$ are the standard coordinates on $T^*\mathbb{H}$ we get

$$\begin{aligned} H((x, y), (p_x, p_y)) \;&=\; \frac{1}{2}\,\|p - \eta\|_z^2 \;=\; \frac{y^2}{2}\left[\left(p_x - \frac{1}{y}\right)^2 + p_y^2\right] \;= \\ &=\; \frac{y^2}{2}\left(p_x^2 + \frac{1}{y^2} - \frac{2p_x}{y} + p_y^2\right) \;=\; \frac{1}{2}\,y^2\,(p_x^2 + p_y^2) - y\,p_x + \frac{1}{2} \;= \\ &=\; \frac{1}{2}\|p\|_z^2 - y\,p_x + \frac{1}{2} \end{aligned}$$

We can also write down the Hamiltonian system in coordinates as

$$\begin{cases} \dot{z} = \; y^2\,p - y \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \\ \dot{p} = \left(p_x - \dfrac{1}{y}\|p\|_z^2\right) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \end{cases}$$

and the Hamilton-Jacobi equation $H(z, du(z)) = k$ as

$$\frac{y^2}{2}\,|\nabla u|^2 - y \cdot \frac{\partial u}{\partial x} \;=\; k - \frac{1}{2}\,.$$

5.2 Dynamics for low energy levels

Let L be the Lagrangian in (5.1); since L is autonomous (i.e. not time dependent) the energy function $E(z, v)$ is a prime integral. Hence we have the equality

$$\dot{x}^2 + \dot{y}^2 = 2cy^2$$

where c represents the energy value; hereafter we suppose $c < \frac{1}{2}$. Moreover since L does not depend explicitly on x the associated momentum $p_x = L_{v_x}$ is also constant along the motion; if c' denotes the (constant) momentum p_x, after a simple computation we get

$$\dot{x} = (c'y - 1)\,y\,.$$

We use the notation $\dot{x}, \dot{y}$ instead of v_x, v_y because we want to point out the time dependence of the solutions. Replacing the value of $\dot{x}$ in the conservation of energy we get

$$\dot{y}^2 = y^2\left[2c - (c'y - 1)^2\right]; \tag{5.7}$$

this equation makes sense if and only if the righthand-side is ≥ 0, that is if and only if $(c'y-1)^2 \leq 2c$. This implies that a solution $y(t)$ must satisfy

$$-\sqrt{2c} \;\leq\; c'y - 1 \leq \sqrt{2c}\,;$$

in particular, since $\sqrt{2c} < 1$ and $y > 0$, we get that c' must be strictly positive and

$$\frac{1-\sqrt{2c}}{c'} \;\leq\; y \;\leq\; \frac{1+\sqrt{2c}}{c'}\,. \tag{5.8}$$

Therefore a solution $y(t)$ of the system is necessary bounded away from 0 and ∞ by two constants depending on the values of the prime integrals. Recall that the solutions must satisfy

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{v_x}{y^2}+\dfrac{1}{y}\right)=0\,; \\ \\ \dfrac{d}{dt}\left(\dfrac{v_y}{y^2}\right)=-\dfrac{1}{y^3}\left(v_x^2+v_y^2\right)-\dfrac{v_x}{y^2}\,; \end{cases} \tag{5.9}$$

Observe that the first Euler-Lagrange equation is the conservation of the momentum p_x; using the conservation of energy and of the momentum p_x in the second equation we can write

$$\frac{d}{dt}\left(\frac{\dot y}{y^2}\right)=-\frac{2c}{y}-\frac{\dot x}{y^2}=\frac{1-2c}{y}-c'\,.$$

If we denote by $u=\frac{1}{y}$, the equation above can be rewritten as

$$\ddot u+(1-2c)u-c'=0\,; \tag{5.10}$$

since $1-2c>0$, the general solution of (5.10) is

$$u(t)=A\cdot\sin\left(\sqrt{1-2c}\cdot t\right)+B\cdot\cos\left(\sqrt{1-2c}\cdot t\right)+\frac{c'}{1-2c}\,.$$

Equation (5.7) implies a condition on the coefficients A, B in order to get an effective solution $y(t)$ of the Euler-Lagrange equation; more precisely, using the u functions we can rewrite (5.7) as

$$\dot u^2=(2c-1)u^2+2c'u-(c')^2\,. \tag{5.11}$$

Now $\dot u(t)=\sqrt{1-2c}\cdot\left[A\cdot\cos\left(\sqrt{1-2c}\cdot t\right)+B\cdot\sin\left(\sqrt{1-2c}\cdot t\right)\right]$ and hence we get

$$\begin{aligned} &(1-2c)\Big[A^2\cdot\cos^2(\alpha t)+B^2\cdot\sin^2(\alpha t)-2AB\cdot\sin(\alpha t)\cos(\alpha t)\Big]= \\ =\;&(2c-1)\Big[A^2\cdot\sin^2(\alpha t)+B^2\cdot\cos^2(\alpha t)+\frac{(c')^2}{(1-2c)^2}+2AB\cdot\sin(\alpha t)\cos(\alpha t)+ \\ +\;&\frac{2Ac'}{1-2c}\cdot\sin(\alpha t)+\frac{2Bc'}{1-2c}\cdot\cos(\alpha t)\Big]+2Ac'\cdot\sin(\alpha t)+2Bc'\cdot\cos(\alpha t)+\frac{2(c')^2}{1-2c}-(c')^2 \end{aligned}$$

where $\alpha=\sqrt{1-2c}$. This horrific equality after simple computations can be rewritten as

$$A^2+B^2=\frac{2c(c')^2}{(1-2c)^2}>0$$

since $c'>0$ and $0<c<\frac{1}{2}$. In particular we get that A,B are not both zero; therefore $u(t)$ (and hence also $y(t)$) is non constant and periodic with period depending only on the energy c. Another way to prove that the Euler-Lagrange equations have no constant y solutions is the following; suppose that y is a constant solution, then the second Euler-Lagrange equation becomes

$$0=\frac{\dot x^2}{y^3}-\frac{\dot x}{y^2}=-\frac{\dot x}{y}\left(\frac{\dot x}{y^2}+\frac{1}{y}\right)=-c'\cdot\frac{\dot x}{y^2}\,;$$

since $\dot x$ can not be zero, it must be $c'=0$, which is in contradiction with $c'>0$.

Proposition 5.2.1. *Let L be the Lagrangian in (5.1) and let $c<\frac{1}{2}$ be a subcritical energy value; then the orbits $t\longmapsto y(t)$ with $E\equiv c$ are periodic and bounded away from 0 and ∞ by two constant depending on the values of the energy and of p_x.*

5.3 Dynamics for supercritical energy values

In this section we want to study the dynamics for energy levels $E \equiv k$ with $k > \frac{1}{2} = c(L)$; it is a general fact, and we have proved it in section 3.3, that the Euler-Lagrange flow on a supercritical energy level is the reparametrization of the geodesic flow for a suitable Finsler metric on its unit tangent bundle. Moreover if $u \in C^\infty(M, \mathbb{R})$ is a subsolution for the Hamilton-Jacobi equation, that is u satisfies the inequality

$$H(z, du(z)) < k\,,$$

then one can take the Finsler Lagrangian F^2 defined on the energy level $E(z,v) = k$ by

$$F(z,v) \;=\; L(z,v) + k - du(z)[v]$$

and extended to be positively homogeneous on the whole tangent bundle TM. Observe that, since by the Fenchel transform

$$H(z,0) = \max_{v \in T_zM} \Big[< 0, v >_z - L(z,v) \Big] = - \min_{v \in T_zM} L(z,v)\,,$$

if $k > -\inf L(z,v)$ then one can choose $u \equiv 0$. Since in our case the critical value $c(L) = \frac{1}{2}$ coincides with the minimum of the Lagrangian among TM (see next section), for any supercritical energy level $E \equiv k$ we can choose $u \equiv 0$ and the Finsler lagrangian F^2 as

$$F(z,v) \;=\; L(z,v) + k \;=\; 2k + \frac{v_x}{y} \tag{5.12}$$

on the energy level $E = k$. In the last equality we have used the fact that $L = E + \eta$, indeed our lagrangian is a magnetic lagrangian and for any magnetic lagrangian the energy is given by

$$E(z,v) \;=\; \frac{1}{2}\|v\|_z^2 + U(z) \;=\; L(z,v) - \eta_z(v) + 2U(z)\,.$$

Now we want to extend F to a Finsler metric on M by requirring that F is positive homogeneous and coincides with (5.12) on the energy level $E \equiv k$; however before we want to observe that F as defined above is effectively always positive. One can check that a necessary condition for a point (z,v) to belong to the energy level $E = k$ is that $|v_x| \le \sqrt{2k} \cdot y$, indeed

$$\frac{v_x^2}{2y^2} \;\le\; \frac{v_x^2 + v_y^2}{2y^2} \;=\; E(z,v) \;=\; k$$

implies that $v_x^2 \le 2ky^2$ and then the condition above clearly follows. Thus we get

$$F(z,v) \;=\; 2k + \frac{v_x}{y} \;\ge\; 2k - \sqrt{2k} \;>\; 0$$

because, by hypothesys, $k > \frac{1}{2}$. Notice that the critical value $c(L) = \frac{1}{2}$ is the infimum of the real numbers k such that equation (5.12) defines effectively a Finsler metric on $\mathbb{H}$; of course, as we have just proved, equation (5.12) for $k = c(L)$ can not define a Finsler metric because

$$F((x,y),(-y,0)) \;=\; 1 - 1 \;=\; 0\,.$$

Since we have proved that F defined above is effectively always positive for $k > \frac{1}{2}$, we can extend it by homogeneity to a Finsler metric on $\mathbb{H}$; more precisely given a point $(z,w) \in T\mathbb{H}$ there exists only one $\lambda \in \mathbb{R}$, $\lambda > 0$ such that $(z, \lambda w) \in E^{-1}(k)$, indeed if we denote by k' the energy of the point (z,w) and by w_x, w_y the coordinates of w, then

$$k \;=\; E(z,\lambda w) \;=\; \frac{\lambda^2 w_x^2 + \lambda^2 w_y^2}{2y^2} \;=\; \lambda^2 \cdot \frac{w_x^2 + w_y^2}{2y^2} \;=\; \lambda^2 E(z,w) \;=\; \lambda^2 k'$$

and hence $\lambda = \sqrt{k/k'}$. Therefore if $\mu = 1/\lambda$ and $v \in T_zM$ is such that $w = \mu v$, then the point (z,v) belongs to the energy level $E \equiv k$ where the Finsler metric is already defined and this, together with the positive homogeneity, allows us to define $F(z,w)$ as follows

$$
\begin{aligned}
F(z,w) &= F(z,\mu v) \;=\; \mu\, F(z,v) \;=\; \mu\left(2k+\frac{v_x}{y}\right) \;=\\
&= \sqrt{k'/k}\cdot 2k+\frac{\mu v_x}{y} \;=\; 2\sqrt{kk'}+\frac{w_x}{y} \;=\; \sqrt{2k\cdot\|w\|_z^2}+\frac{w_x}{y} \;=\\
&= \sqrt{2k}\cdot\|w\|_z+\eta_z(w)\,.
\end{aligned}
$$

Observe that the Finsler metric $F(z,w)$ just defined is effectively positive homogeneous and if (z,w) belongs to the energy level $E=k$ then

$$
F(z,w) \;=\; \sqrt{2k}\cdot\|w\|_z+\eta_z(w) \;=\; 2k+\eta_z(w) \;=\; 2k+\frac{w_x}{y}
$$

which is precisely equation (5.12). We have just proved the following

Proposition 5.3.1. *Let $L: T\mathbb{H}\longrightarrow\mathbb{R}$ be the Lagrangian defined in (5.1) and let $k>\frac{1}{2}$ be a supercritical energy value; then the Euler-Lagrange flow on the energy level $E=k$ is a reparametrization of the geodesic flow on its unit tangent bundle with respect to the Finsler Lagrangian*

$$
F(z,v)^2 \;=\; \frac{1}{2k}\left(\sqrt{2k}\cdot\|v\|_z+\eta_z(v)\right)^2\,. \tag{5.13}
$$

The choice of the constant $\frac{1}{2k}$ in equation (5.13) is motivated from the fact that the limit for $k\to\infty$ of F^2 is exactly the riemannian Lagrangian $L(z,v)=\frac{1}{2}\|v\|_z^2$, indeed

$$
F(z,v)^2 \;=\; \frac{1}{2k}\left(\sqrt{2k}\cdot\|v\|_z+\eta_z(v)\right)^2 \;=\; \frac{1}{2}\|v\|_z^2+\frac{2}{\sqrt{2k}}\cdot\eta_z(v)\,\|v\|_z+\frac{1}{2k}\cdot\eta_z(v)^2 \;\longrightarrow\; \frac{1}{2}\|v\|_z^2\,.
$$

Therefore, for high energy value the Finsler metric converges to the hyperbolic metric.

5.4 The Hamilton-Jacobi equation: a geometric approach

Consider the Hamilton-Jacobi equation

$$
y^2\,|\nabla u|^2-2y\cdot\frac{\partial u}{\partial x} \;=\; 2k-1 \tag{5.14}
$$

associated to the Hamiltonian (5.6). We already know from what we have proved in chapter 3 that any smooth solution $u:\mathbb{H}\longrightarrow\mathbb{R}$ of (5.14) corresponds to an invariant exact Lagrangian graph, while smooth subsolutions correspond to exact Lagrangian graphs; moreover there are no (weakly) differentiable subsolutions of (5.14) for any $k<\frac{1}{2}=c(L)$ while for any $k>\frac{1}{2}$ there is always a smooth subsolution. This allows to see the critical value as the infimum of the values $k\in\mathbb{R}$ such that the sublevel $\{H<k\}$ contains an exact Lagrangian graph. Here we are interested into compute the (smooth) solutions of the Hamilton-Jacobi equation in the case $k=\frac{1}{2}$; rather than trying to solve the PDE (5.14) with standard analitical methods, we use a more geometric approach which consists into determine the invariant exact Lagrangian graphs. In the case $k=\frac{1}{2}$ the Hamilton-Jacobi equation can be rewritten as

$$
y\,|\nabla u|^2-2\cdot\frac{\partial u}{\partial x} \;=\; 0
$$

and hence we get immediately that the constant functions $u\equiv c$ are solutions. This implies that $\mathbb{H}\times\{0\}$ is an exact invariant Lagrangian graph in $T^*\mathbb{H}$; another way to see this fact is the following: let L be the Lagrangian in (5.4) and denote by (v_x,v_y) the standard coordinates on $T_z\mathbb{H}$; therefore, if we suppose $z=(x,y)\in\mathbb{H}$ fixed, we have

$$
L(z,v) \;=\; \frac{1}{2y^2}(v_x^2+v_y^2)+\frac{v_x}{y} \;\geq\; \frac{v_x^2}{2y^2}+\frac{v_x}{y}
$$

where the righthand-side attains its minimum $-\frac{1}{2}$ when $v_x=-y$. Thus the Lagrangian L attains its minimum among $T_z\mathbb{H}$ at $(-y,0)$ and

$$L((x,y),(-y,0)) \;=\; -\frac{1}{2}\,.$$

Since the minimum of the Lagrangian L among any tangent space $T_z\mathbb{H}$ is the same (i.e. does not depend on z) and equal to $\frac{1}{2}$, applying the Fenchel transform we get

$$H(z,0) \;=\; \max_{v\in T_z\mathbb{H}}\Big[<0,v>_z - L(z,v)\Big] \;=-\min_{v\in T_z\mathbb{H}} L(z,v) \;=\; \frac{1}{2}\,.$$

The calculation above shows also that for any $k>\frac{1}{2}$ the constant functions are (smooth) subsolutions for the Hamilton-Jacobi equation $H(z,du(z))=k$. This argument clearly also applies for any general Lagrangian $L:TM\longrightarrow\mathbb{R}$ and assures that for any $k>-\inf L(x,v)$ the constant functions are smooth subsolution of the Hamilton-Jacobi equation; in our case we get something more (that is the constant functions are effectively solutions) which may not hold in the general case. After this simple remark we may proceed with a more general and interesting argument; as already mentioned, we try to compute smooth solutions of the Hamilton-Jacobi equation for $k=\frac{1}{2}$ using a geometric approach which consists into determine the shape of any invariant (with respect to the Hamiltonian flow associated to the horocycle flow through the Legendre transform) graph in $T^*\mathbb{H}$ and into investigate which of these invariant graph are also exact Lagrangian. Observe that if an invariant graph is Lagrangian, then it is also exact; indeed since $\mathbb{H}$ is simply connected any closed 1-form in $\mathbb{H}$ is exact and we have observed in chapter 3 that a graph G_η is lagrangian if and only if the 1-form η is closed. Moreover any invariant graph in $T^*\mathbb{H}$ corresponds via the Legendre transform

$$\mathcal{L}:T\mathbb{H}\longrightarrow T^*\mathbb{H}\,,\quad (z,v)\longmapsto\left(z\;,\left(\frac{v_x}{y^2}+\frac{1}{y}\;,\;\frac{v_y}{y}\right)\right) \tag{5.15}$$

to an invariant graph in $T\mathbb{H}$ (with respect to the Euler-Lagrange flow); therefore it suffices to determine all the invariant graphs in $T\mathbb{H}$ and then see which of them correspond to graphs in $T^*\mathbb{H}$ defined by a closed 1-form. The first step is to show which shape must have a foliation of $\mathbb{H}$; recall that the Euler-Lagrange flow at the energy level $\frac{1}{2}$ is the horocycle flow, hence the images of the solutions are the horospheres. So let $\zeta:\mathbb{H}\longrightarrow T\mathbb{H}$ be a vector field such that the graph

$$\Gamma_\zeta \;=\; \{(z,\zeta(z))\mid z\in\mathbb{H}\}$$

is invariant and contained in the energy level $\{E\equiv\frac{1}{2}\}$; since the energy is of the form $E(z,v)=\frac{1}{2}\,\|v\|_z^2$ we get immediately that $\zeta(z)\neq 0$ for any $z\in\mathbb{H}$ and hence $\int\zeta$ defines an oriented foliation of $\mathbb{H}$, which is composed by horospheres since ζ is invariant.

Lemma 5.4.1. *If $\mathcal{F}$ is a smooth foliation of $\mathbb{H}$ then it is composed by horospheres tangent to a same suitable point $a\in\partial\mathbb{H}$; in particular if $a=\infty$ then the foliation is made by horizontal lines.*

Proof. Fix a point $a\in\partial\mathbb{H}$; it is clear that the horospheres tangent to $\partial\mathbb{H}$ in a form a foliation of $\mathbb{H}$, indeed for any point $x\in\mathbb{H}$ there is only one horosphere through p and tangent to $\partial\mathbb{H}$ in a. We denote by $\mathcal{F}_a$ such foliation; now suppose that $\mathcal{F}$ is another foliation of $\mathbb{H}$ which is not of the form $\mathcal{F}_a$ for any $a\in\partial\mathbb{H}$. Observe that if $\mathcal{F}$ contains a horosphere h tangent to $\partial\mathbb{H}$ in a point a then $\mathcal{F}$ contains all the horospheres tangent to $\partial\mathbb{H}$ in a and contained in the circle defined by h; indeed if z is an interior point of the circle whose boundary is h then the only horosphere through z that does not intersect h is the horosphere tangent to $\partial\mathbb{H}$ in a. So if $\mathcal{F}$ is not of the form $\mathcal{F}_a$ then there exists a point $b\in\partial\mathbb{H}$ such that $\mathcal{F}$ contains a family $\{h_r\}_{r<R}$ of horospheres with (Euclidean) radius $<R$ tangent to $\partial\mathbb{H}$ in b and no horospheres tangent in b to $\partial\mathbb{H}$ with Euclidean radius $r>R$. So if $z=(b,y)$ with $y>R$ then the horosphere through z must be the horizontal line l_y, indeed the only other horosphere through z that does not intersect the family $\{h_r\}_{r>R}$ is the horosphere h_y, which does not belong to $\mathcal{F}$. Now if we consider a point $z\in\mathbb{H}\setminus h_R$ sufficiently close to (b,R) as in the figure above we get that there are no horospheres through z with no intersections with the family $\{h_r\}_{r<R}$ and the horizontal lines $\{l_y\}_{y>R}$ and hence $\mathcal{F}$ is not a foliation. $\square$

Observe that the foliation $\mathcal{F}_\infty$ is composed by horizontal lines while the foliation $\mathcal{F}_0$ is composed by horospheres tangent to $\partial\mathbb{H}$ in the origin; since a foliation determines the vector field ζ uniquely up to a multiplicative factor, we get that any invariant graph is of the form

$$\alpha \cdot \Gamma_a \;=\; \{(z, \alpha\, v_z^a) \mid z \in \mathbb{H}\}$$

where α is a positive number (since the foliation is oriented) and v_z^a is the unit tangent vector at z to the horosphere through z and tangent to $\partial\mathbb{H}$ in a. Now, since the graph is contained in the energy level $E = \frac{1}{2}$, it must be $\alpha = 1$ and hence we can write simply Γ_a instead of $1 \cdot \Gamma_a$; in the particular case $a = \infty$ we get

$$\Gamma_\infty \;=\; \{((x,y),(-y,0)) \mid (x,y) \in \mathbb{H}\}$$

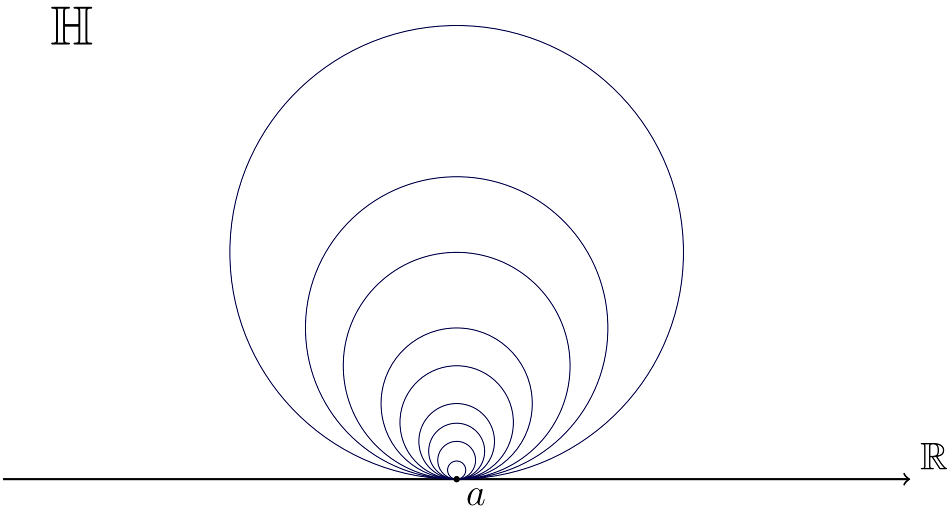


Figure 5.2 *The smooth foliations of $\mathbb{H}$ made by horospheres.*

with the corresponding invariant graph $\mathcal{L}(\Gamma_\infty) = \mathbb{H} \times \{0\}$ that is clearly exact Lagrangian; we have obtained in another way that the constant functions are solutions for the Hamilton-Jacobi equation. Now consider any $a \in \partial\mathbb{H} \setminus \{\infty\}$; the invariant graph Γ_a is given by

$$\Gamma_a \;=\; \left\{\left(\eta_b^a(t), \dot\eta_b^a(t)\right) \;\middle|\; t \in \mathbb{R}, b \in \mathbb{R}\right\}$$

where $\eta_b^a(t)$ is, as in equation (4.9), the arclength parametrisation of the horospheres tangent to $\partial\mathbb{H}$ at the point a. In particular the unit tangent vector $\dot\eta_b^a(t)$ at any point $\eta_b^a(t)$ is

$$\dot\eta_b^a(t) \;=\; \frac{1}{b^2(1+t^2)^2}\left(b(1+t^2) - 2bt^2, -2bt\right) \;=\; \frac{1}{b(1+t^2)^2}\left(1-t^2, -2t\right). \tag{5.16}$$

We want now to write Γ_a with respect to the standard coordinates of the hyperbolic plane. Imposing that the first coordinate of $\eta_b^a(t)$ is equal to x and that the second is equal to y we get the system

$$\begin{cases} x \;=\; a + \dfrac{t}{b(1+t^2)}\,; \\[2ex] y \;=\; \dfrac{1}{b(1+t^2)}\,; \end{cases}$$

from which we obtain

$$t \;=\; \frac{x-a}{y}, \qquad\qquad b \;=\; \frac{1}{y(1+t^2)} \;=\; \frac{1}{y\left(1+\frac{(x-a)^2}{y^2}\right)} \;=\; \frac{y}{(x-a)^2+y^2}\,.$$

Now it is easy to see that in the standard coordinate system $\dot\eta_b^a(t)$ can be written as

$$\dot\eta_b^a(t) \;=\; b\left(y^2 - x^2, -2xy\right) \;=\; \frac{y}{(x-a)^2+y^2}\left(y^2-x^2, -2xy\right)$$

and hence

$$\Gamma_a \;=\; \left\{ \left((x,y)\,, \left(\frac{y}{(x-a)^2+y^2}\,\big(y^2-x^2,-2xy\big) \right) \right) \;\Big|\; (x,y)\in\mathbb{H} \right\}. \tag{5.17}$$

Therefore the corresponding graph $\Sigma_a \subseteq T^*\mathbb{H}$, given by the Legendre tranform $\mathcal{L}(\Gamma_a)$ of Γ_a, is of the form $\Sigma_a = G_{\omega_a} = \{(z,\omega_a(z)) \mid z\in\mathbb{H}\}$ where ω_a is the 1-form

$$\omega_a(x,y) \;=\; \frac{2y}{(x-a)^2+y^2}\,dx - \frac{2(x-a)}{(x-a)^2+y^2}\,dy\,.$$

It is easy to check that the 1-form ω_a is closed (and hence exact), since the partial derivatives of the first factor of ω_a with respect to y and of the second factor of ω_a with respect to x are equal

$$\frac{\partial(\omega_a)_x}{\partial y} \;=\; \frac{2[(x-a)^2+y^2]-4y^2}{[(x-a)^2+y^2]^2} \;=\; \frac{2(x-a)^2-2y^2}{[(x-a)^2+y^2]^2} \;=\; \frac{\partial(\omega_a)_y}{\partial x}\,.$$

Thus there exists a smooth function $u_a \in C^\infty(\mathbb{H})$ such that $du_a=\omega_a$; such a primitive is obtained imposing the conditions on the partial derivatives of u_a. The relation $\frac{\partial u_a}{\partial x} = (\omega_a)_x$ implies that

$$u_a(x,y) \;=\; \int \frac{2y}{(x-a)^2+y^2}\,dx \;=\; \int \frac{2}{1+\zeta^2}\,d\zeta \;=\; 2\cdot\arctan\zeta + g(y)$$

where we have made the change of variable $\zeta = \frac{x-a}{y}$, while $\frac{\partial u_a}{\partial y} = (\omega_a)_y$ implies that

$$\frac{\partial u_a}{\partial y} \;=\; \frac{-2(x-a)}{(x-a)^2+y^2} + g'(y);$$

therefore we get that g is constant and a primitive of ω_a is the function

$$u_a(x,y) \;=\; 2\cdot\arctan\left(\frac{x-a}{y}\right). \tag{5.18}$$

Hence G_{du_a} is an exact invariant Lagrangian graph and the function u_a is a solution of the Hamilton-Jacobi equation. This sort of invariance of Hamilton-Jacobi solutions under translations of the first variable $x \longmapsto x+a$ could also be a priori guessed since the Lagrangian does not depend explicitly on x and hence if $u(x,y)$ is a solution of the Hamilton-Jacobi equation then it is reasonable to believe that $u_a(x,y)=u(x-a,y)$ is also solution for any $a\in\mathbb{R}$. Observe that the solutions obtained with this method are the only solutions of the Hamilton-Jacobi equation; indeed any solution corresponds to a (unique) invariant Lagrangian graph, which corresponds to a unique invariant graph in $T\mathbb{H}$. Furthermore any invariant graph in $T^*\mathbb{H}$ is exact Lagrangian and hence it corresponds to a (unique) solution of the Hamilton-Jacobi equation; this represents in a certain sense the viceversa of the Hamilton-Jacobi theorem in chapter 3, which states that a Lagrangian manifold (in particular a Lagrangian graph) contained in an energy level is invariant. Summarizing we have proved

Theorem 5.4.2. *Let H be the Hamiltonian in (5.6) and let*

$$y\,|\nabla u|^2 - 2\cdot\frac{\partial u}{\partial x} \;=\; 0 \tag{5.19}$$

be the associated Hamilton-Jacobi equation at level $k=\frac{1}{2}$; then the following hold:

1. *If u is a smooth solution of (5.19) then u is constant or $u=u_a$ for some $a\in\partial\mathbb{H}$.*
2. *If ω is a 1-form on $\mathbb{H}$ such that $G_\omega \subseteq \{H\equiv\frac{1}{2}\}$ is invariant, then ω is exact.*

We show now another method, that will be usefull also in the next section, to compute the functions u_a; if in the method above we started from the arc-length parametrization of the horizontal horospheres and, through the composition with a suitable isometry, we obtained the arc-length parametrization of the horospheres tangent to $\partial\mathbb{H}$ in a fixed point a, now we are going to parametrize the invariant graphs Γ_a, $a\in\mathbb{R}$ in an easier coordinate system. The construction is analogous for any Γ_a (it suffices to translate the first coordinate of (5.20) by a factor a), so we can consider the invariant graph Γ_0 composed by the horospheres tangent to $\partial\mathbb{H}$ at the origin and introduce the coordinate system (r, ϑ) for the hyperbolic plane

$$\begin{cases} x = -r \sin\vartheta\,; \\ \\ y = r\,(1-\cos\vartheta)\,; \end{cases} \tag{5.20}$$

where $r > 0$ and $\vartheta \in (0, \pi)$. If we fix r then the curve $\gamma_r(\vartheta) = r(-\sin\vartheta, 1-\cos\vartheta)$ parametrizes the horosphere of Euclidean radius r and tangent to $\partial\mathbb{H}$ at the origin, as shown in figure 5.3; observe that the parametrization $\gamma_r(\vartheta)$ is in general not by arc-length with respect neither to the Euclidean metric nor to the hyperbolic metric of $\mathbb{H}$. However the tangent vector at a point $\vartheta \in (0, \pi)$ is given by $\gamma_r'(\vartheta) = r(-\cos\vartheta, \sin\vartheta)$, so that the (hyperbolic) unit tangent vector is

$$v_r(\vartheta) \;=\; r(1-\cos\vartheta)\big(-\cos\vartheta, \sin\vartheta\big)\,.$$

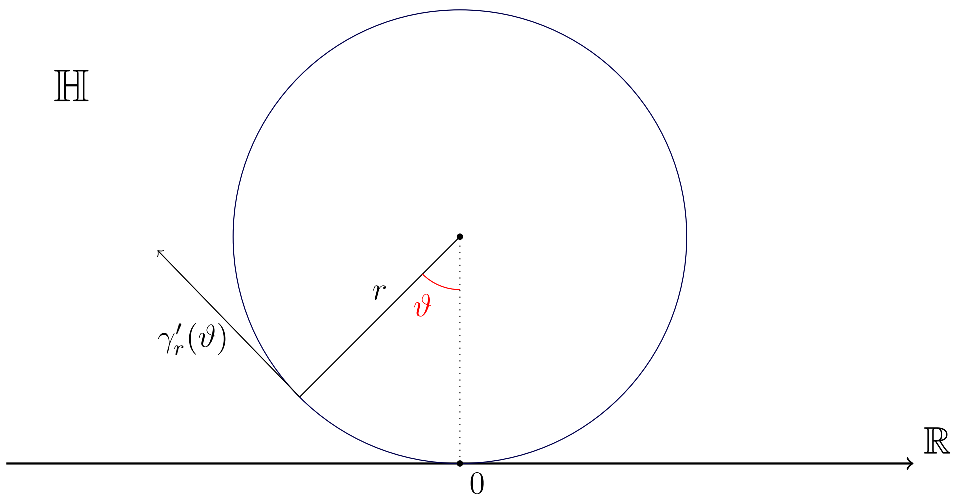


Figure 5.3 *The curves $\gamma_r : (0, \pi) \longrightarrow \mathbb{H}$ parametrize the horospheres tangent to $\partial\mathbb{H}$ at 0.*

By equation (5.20) we get immediately that

$$r \;=\; \frac{x^2+y^2}{2y}\,, \qquad \sin\vartheta \;=\; \frac{-2xy}{x^2+y^2}\,, \qquad \cos\vartheta \;=\; \frac{x^2-y^2}{x^2+y^2}$$

and hence

$$\begin{aligned} \Gamma_0 \;&=\; \Big\{\Big(r(-\sin\vartheta, 1-\cos\vartheta),\; r(1-\cos\vartheta)\big(-\cos\vartheta,\; \sin\vartheta\big)\Big) \;\Big|\; \vartheta \in (0, 2\pi),\, r > 0\Big\} \;= \\ &=\; \Big\{\Big((x,y)\;, \frac{y}{x^2+y^2}(y^2-x^2, -2xy)\Big) \;\Big|\; (x,y) \in \mathbb{H}\Big\} \end{aligned}$$

which is (5.17); now, exactly as we have done before, we get the thesys mapping Γ_0 through the Legendre trasform to the corresponding invariant graph in $T^*\mathbb{H}$ and proving that this graph is exact.

5.4.1 Simple remarks in the case $k > \frac{1}{2}$

Consider the Hamilton-Jacobi equation (5.14) and suppose now that $k > \frac{1}{2}$; we are interested into compute simple solutions of this equation, as for istance solutions of the form $u(x, y) = f(x) \cdot g(y)$ for suitable smooth functions $f \in C^\infty(\mathbb{R})$, $g \in C^\infty(0, +\infty)$. For such solutions we can rewrite the Hamilton-Jacobi equation as

$$y^2\left(f'(x)^2 + g'(y)^2\right) - 2y\,f'(x) \;=\; 2k-1\,.$$

Hence we get that the equation

$$f'(x)^2 - \frac{2}{y}\, f'(x) \;=\; -g'(y)^2 + \frac{2k-1}{y^2} \tag{5.21}$$

in which the righthand-side depends only on y while the lefthand-side depends on both; therefore the lefthand-side must be constant

$$f'(x)^2 \;=\; \frac{2}{y}\, f'(x) + c$$

but in this new equation the LHS depends only on x while the RHS depends on both x, y and hence the only possibility is that f is constant and $c = 0$. By (5.21) we get that

$$g'(y)^2 \;=\; \frac{2k-1}{y^2}\,; \tag{5.22}$$

Since $k > \frac{1}{2}$, we get that a solution g of (5.22) effectively exists. Observe that, in agreement with what we have proved in the previous section, there are no smooth non constant solutions depending separately on x and y in the case $k = \frac{1}{2}$. Equation (5.22) for $k > \frac{1}{2}$ has solution $g(y) = \sqrt{2k-1} \cdot \log y + \mu$ so that

$$u(x, y) = \lambda \sqrt{2k-1} \cdot \log y + \mu\,, \qquad \lambda, \mu \in \mathbb{R}$$

can be solutions of the Hamilton-Jacobi equation. Now for such functions u we have

$$H(z, du(z)) \;=\; y^2\, \lambda^2 \cdot \frac{2k-1}{y^2} \;=\; \lambda^2(2k-1)$$

and hence it must be $\lambda = \pm 1$. We have just proved the following

Lemma 5.4.3. *Any smooth solution $u : \mathbb{H} \longrightarrow \mathbb{R}$ of the Hamilton-Jacobi equation (5.14) for $k > \frac{1}{2}$ depending separately on x, y is of the form*

$$g(y) = \pm\sqrt{2k-1} \cdot \log y + \mu\,.$$

5.5 Smooth Hamilton-Jacobi solutions for the geodesic flow

Hear we want to compute, using the method explained in the last section, some particular smooth solutions $f : \mathbb{H} \longrightarrow \mathbb{R}$ of the Hamilton-Jacobi equation associated to the geodesic flow of $\mathbb{H}$; recall that the geodesics of the hyperbolic plane are the semicircles orthogonal to $\mathbb{R}$ and the vertical lines parametrized by arc-length. The geodesic flow can be seen (section 1) as the Euler-Lagrange flow for the riemannian Lagrangian

$$L(z, v) \;=\; \frac{1}{2}\|v\|_z^2$$

where $\|\cdot\|_z$ denotes the norm induced by the hyperbolic metric. Since the Euler-Lagrange flow at the energy level $E \equiv k$ is the geodesic flow up to a uniform change of speed, we can study the problem at the energy level $\frac{1}{2}$ (in which we have the geodesics parametrized by arclength); the Hamilton-Jacobi equation $H(z, du(z)) = \frac{1}{2}$ has the form

$$H(z, du(z)) \;=\; \frac{1}{2}\,\|du\|_z^2 \;=\; y^2\,\|du\|_E^2 \;=\; \frac{1}{2}$$

where here $\|\cdot\|_z$ denotes the dual norm on $T^*\mathbb{H}$ induced by the norm on $T\mathbb{H}$ and $\|\cdot\|_E$ denotes the Euclidean norm. So a smooth function u is solution of the Hamilton-Jacobi equation if and only if

$$|\nabla u|^2 \;=\; \frac{1}{y^2}\,. \tag{5.23}$$

However we shall observe that the method in section 5.4 does not find all the solutions of (5.23) because we consider just particular foliations of $\mathbb{H}$, that is the foliations composed by geodesics ending in the same point $a \in \partial\mathbb{H}$. In the horocycle flow case (lemma 5.4.1) the foliations made by horospheres ending at the same point were the only possible; here this is no longer the case, as we will see at the end of this section constructing other examples of foliations composed by geodesics with the same center $a \in \partial\mathbb{H}$ (it is also possible to get other foliations letting the center free); these foliations define other solutions of the Hamilton-Jacobi equation. So let us start with the foliation of $\mathbb{H}$ made by the vertical lines parametrized by arc-length, that is $\gamma_b(t) = (b, e^t)$ for any $b \in \mathbb{R}$, $t \in \mathbb{R}$, and consider the invariant graph

$$\Gamma = \left\{ \left((b, e^t), (0, e^t) \right) \;\middle|\; b \in \mathbb{R}\,,\ t \in \mathbb{R} \right\} = \left\{ \left((x, y), (0, y) \right) \;\middle|\; (x, y) \in \mathbb{H} \right\}.$$

The corresponding invariant graph in $T^*\mathbb{H}$ through the Legendre transform is

$$\Sigma \;=\; \left\{ \left((x, y), \left(0\,, \frac{1}{y} \right) \right) \;\middle|\; (x, y) \in \mathbb{H} \right\}$$

and it is clear that Σ is the graph of the form du, where $u = \log y$; hence u is a solution for (5.23). Since the Lagrangian is simmetric, i.e. $L(z, v) = L(z, -v)$, the Euler-Lagrange flow is reversible on the time; hence there is another invariant graph associated to the foliation in vertical lines. Such graph is obtained considering the opposite arclength parametrization of the vertical geodesics, that is $\gamma'_b(t) = (b, e^{-t})$, and yields to the solution $u = -\log y$ of the Hamilton-Jacobi equation. This fact could be also deduced from the quadratic dependence on (the derivatives of) u in (5.23) and implies that if we get a solution u of the Hamilton-Jacobi equation then also $-u$ is solution. Now consider the foliation of $\mathbb{H}$ made by geodesics ending in the same point $a \in \mathbb{R}$; the invariant graph associated to this foliation is (up to a time inversion) given by

$$\Gamma_a \;=\; \left\{ \left(\eta_a^b(t), \dot\eta_a^b(t) \right) \;\middle|\; b, t \in \mathbb{R} \right\}$$

where $\eta_a^b(t)$ is the arclength parametrisation of the geodesic with center in a given by the formula (4.4). In particular the unit tangent vector at any point is

$$\dot\eta_a^b(t) \;=\; \frac{1}{(b^2 + e^{2t})^2} \left(2be^{2t}, e^t(b^2 + e^{2t}) - 2e^t e^{2t} \right) \;=\; \frac{e^t}{(b^2 + e^{2t})^2} \left(2e^t b, b^2 - e^{2t} \right)$$

In standard coordinate system

$$\begin{cases} x \;=\; a - \dfrac{b}{b^2 + e^{2t}}\,; \\[2ex] y \;=\; \dfrac{e^t}{b^2 + e^{2t}}\,; \end{cases}$$

we have $b = -e^t \cdot \dfrac{x-a}{y}$ and hence $e^t = \dfrac{y}{(x-a)^2 + y^2}$; thus

$$\Gamma_a \;=\; \left\{ \left((x, y)\,, \frac{y}{(x-a)^2 + y^2} \left(-2(x-a)y, (x-a)^2 - y^2 \right) \right) \;\middle|\; (x, y) \in \mathbb{H} \right\}.$$

The corresponding graph in $T^*\mathbb{H}$ through the Legendre transform is

$$\Sigma_a \;=\; \left\{ \left((x, y), \frac{1}{(x-a)^2 + y^2} \left(-2(x-a), \frac{(x-a)^2 - y^2}{y} \right) \right) \;\middle|\; (x_1, x_2) \in \mathbb{H} \right\}$$

that is the graph G_{ω_a} of the 1-form

$$\omega_a(x, y) \;=\; -\,\frac{2(x-a)}{(x-a)^2 + y^2}\, dx \;+\; \frac{(x-a)^2 - y^2}{y[(x-a)^2 + y^2]}\, dy\,.$$

It is just straightforward computation to see that the 1-form ω_a is closed (hence exact) and that a primitive is given by the smooth function $u_a : \mathbb{H} \longrightarrow \mathbb{R}$

$$u_a(x, y) \;=\; \log y - \log((x-a)^2 + y^2)\,. \tag{5.24}$$

Proposition 5.5.1. *The functions $\pm u_a$ defined by (5.24) for any $a \in \partial\mathbb{H}$, $u_\infty(x, y) = \log y$, are solutions for the Hamilton-Jacobi equation (5.23).*

Proof. It is clear that u_∞ is solution; so let u_a as above and compute

$$\begin{aligned} |\nabla u_a(x,y)|^2 &= \frac{4(x-a)^2}{(x-a)^2+y^2} + \frac{[(x-a)^2-y^2]^2}{x^2[(x-a)^2+y^2]^2} = \\ &= \frac{4(x-a)^2y^2+[(x-a)^2-y^2]^2}{y^2[(x-a)^2+y^2]^2} = \frac{[(x-a)^2+y^2]^2}{y^2[(x-a)^2+y^2]^2} = \frac{1}{y^2}\,. \end{aligned}$$

□

Now we want to compute the Hamilton-Jacobi solutions associated to the invariant foliations of $\mathbb{H}$ made by geodesics with the same center $a \in \mathbb{R}$ (recall that geodesics are half-circles or half lines, so the word center makes effectively sense); the argument is analogous for any $a \in \mathbb{R}$, so we may suppose $a = 0$ and then shift the first coordinate by a factor a. Consider the polar coordinate system (r, ϑ) for the hyperbolic plane

$$\begin{cases} x = -r\,\cos\vartheta; \\ \\ y = r\,\sin\vartheta; \end{cases}$$

where $r > 0$ and $\vartheta \in (0, \pi)$; for any r fixed the curve $\gamma_r : \vartheta \longmapsto r\,(-\cos\vartheta, \sin\vartheta)$ parametrizes the geodesic with center in the origin and Euclidean radius r. Although the parametrization is not by arc-length with respect neither to the Euclidean metric nor to the hyperbolic one, we can associate to each point $\gamma_r(\vartheta)$ a (hyperbolic) unit tangent vector

$$v_r(\vartheta) \;=\; r\,\sin\vartheta\,(\sin\vartheta, \cos\vartheta)\,;$$

therefore we get the invariant graph made by the pairs $(\gamma_r(\vartheta), v_r(\vartheta))$, that is

$$\begin{aligned} \Gamma_0 &= \Big\{ \big(r\,(-\cos\vartheta, \sin\vartheta)\,,\; r\,\sin\vartheta\,(\sin\vartheta,\cos\vartheta)\big) \;\Big|\; r>0,\ \vartheta \in (0,\pi) \Big\} = \\ &= \Big\{ \Big((x,y)\,,\; y\left(\frac{y}{r}, -\frac{x}{r}\right)\Big) \;\Big|\; (x,y) \in \mathbb{H} \Big\} = \\ &= \Big\{ \Big((x,y)\,,\; y \cdot \sqrt{x^2+y^2}\left(\frac{y}{x^2+y^2}, -\frac{x}{x^2+y^2}\right)\Big) \;\Big|\; (x,y) \in \mathbb{H} \Big\}\,. \end{aligned}$$

The corresponding graph in $T^*\mathbb{H}$ through the Legendre transform is

$$\Sigma_0 \;=\; \left\{ \left((x,y)\,,\; \left(\frac{1}{\sqrt{x^2+y^2}}, \frac{-x}{y\,\sqrt{x^2+y^2}} \right) \right) \;\Big|\; (x,y) \in \mathbb{H} \right\}$$

which is the graph of the 1-form

$$\omega(x,y) \;=\; \frac{1}{\sqrt{x^2+y^2}}\,dx \;-\; \frac{x}{y\,\sqrt{x^2+y^2}}\,dy\,.$$

The 1-form ω is closed and hence exact, indeed the partial derivative of ω_y with respect to x is

$$\frac{\partial \omega_y}{\partial x} \;=\; \frac{1}{y^2\,(x^2+y^2)} \left(-y\,\sqrt{x^2+y^2} + \frac{x^2 y}{\sqrt{x^2+y^2}} \right) \;=\; \frac{-y\,(x^2+y^2)+x^2y}{y^2\,(x^2+y^2)\,\sqrt{x^2+y^2}} \;=\; \frac{-y}{(x^2+y^2)^{\frac{3}{2}}}$$

and it is easy to see that is equal to the partial derivative of ω_x with respect to y; therefore there exists a smooth function such that $\omega = du$. Such an u satisfies the condition $\frac{\partial u}{\partial x} = \omega_x$ and hence integrating with respect to x we get

$$u(x,y) \;=\; \int \frac{1}{\sqrt{x^2+y^2}}\,dx \;=\; \int \frac{1}{\sqrt{1+\zeta^2}}\,d\zeta \;=\; \operatorname{arcsinh}\zeta + g(y)$$

where we have made the change of variable $\zeta = \frac{x}{y}$; now imposing the condition $\frac{\partial u}{\partial y} = \omega_y$ we get

$$g'(y) - \frac{x}{y\sqrt{x^2+y^2}} \;=\; -\frac{x}{y\sqrt{x^2+y^2}}$$

and hence g is constant. Thus the invariant graph Σ_0 is exact Lagrangian and defined by

$$u(x,u) \;=\; \operatorname{arcsinh}\left(\frac{x}{y}\right)$$

which is a smooth solution of the Hamilton-Jacobi equation. Applying the translation $x \longmapsto x+a$ on the first cordinate we get that for any $a \in \mathbb{R}$ the function

$$u_a(x,y) \;=\; \operatorname{arcsinh}\left(\frac{x-a}{y}\right) \tag{5.25}$$

is a smooth solution of the Hamilton-Jacobi equation. Wa have proved the following

Proposition 5.5.2. *The functions $\pm u_a$ defined in equation (5.25) are smooth solutions of the Hamilton-Jacobi equation (5.23) associated to the geodesic flow of the hyperbolic plane.*

The solutions obtained with these arguments are probably not the unique solutions of (5.23); however, in the examined cases, we started from a certain invariant graph contained in the energy level $E = \frac{1}{2}$ and we proved that there is a (unique) corresponding smooth Hamilton-Jacobi solution. The reader may wonder if this property is true for any invariant graph contained in an energy level; in other words, it is possible that a property analogous to statement 2. in theorem 5.4.2 holds also for the geodesic flow of the hyperbolic plane. Here we do not focus on this problem; however, any suggestion which colud be usefull to solve it is well accepted!

5.6 Horocycle flows for compact quotients of $\mathbb{H}$

Let $M = \Gamma \setminus \mathbb{H}$ be a complete compact hyperbolic manifold and let λ be its standard area form (induced by that of $\mathbb{H}$); denote by $\pi : TM \longrightarrow M$ the canonical projection and define the twisted symplectic form

$$\Omega_\lambda := \; \Omega_0 + \pi^*\lambda$$

where Ω_0 is the symplectic form on TM obtained by pulling back the canonical symplectic form ω_0 on T^*M. Consider now the Hamiltonian flow, also called magnetic flow, defined by the function

$$E(z,v) \;=\; \frac{1}{2}\,\|v\|_z^2 \tag{5.26}$$

and by the form Ω_λ, where as usual $\|\cdot\|_z$ denotes thel norm on TM induced by the hyperbolic metric. As already pointed out in chapter 3, this models the motion of a point under the action of the magnetic field λ; let $p : \mathbb{H} \longrightarrow M$ be the universal covering of M. We already know that the pull back σ of λ through p is exact and a primitive is given by the 1-form $\eta = \frac{dx}{y}$; moreover the solutions of the Euler-Lagrange flow associated to the Lagrangian $L : \mathbb{H} \longrightarrow \mathbb{R}$

$$L(z,v) \;=\; \frac{1}{2}\|v\|_z^2 + \eta_z(v)$$

coincide with the lift to $\mathbb{H}$ of the magnetic geodesics (i.e. the orbits of the magnetic flow) and the energy function of such Lagrangian coincides with the lift to $\mathbb{H}$ of (5.26). In particular the magnetic flow can be obtained projecting on M the Euler-Lagrange flow associated to L; since the horocycle flow on T_1M is defined as the projection of the horocycle flow on $T_1\mathbb{H}$ and this can be seen as the Euler-Lagrange flow of L at energy $\frac{1}{2}$, we get that the horocycle flow on T_1M can be obtained by projecting on T_1M the Euler-Lagrange orbits of L with energy $\frac{1}{2}$. The Euler-Lagrange flow of L can also be viewed through the Legendre transform as the Hamiltonian flow on $T^*\mathbb{H}$ defined by the canonical symplectic form ω_0 and the Hamiltonian

$$H(z,p) \;=\; \frac{1}{2}\,\|p-\eta\|_z^2\,;$$

the critical value of the pair (g, λ), which is defined as

$$c(g,\lambda) \;=\; \inf_{u\in C^\infty(\mathbb{H})}\; \sup_{z\in\mathbb{H}}\; H(z,du(z)) \;=\; \inf_{u\in C^\infty(\mathbb{H})}\; \sup_{z\in\mathbb{H}}\; \frac{1}{2}\,\|du-\eta\|_z^2\,,$$

obviously coincides with the Mañé's critical value $c(L) = \frac{1}{2}$; hence any minimizing measure must be contained in the energy level $\frac{1}{2}$ and must be invariant with respect to the horocycle flow. On the other hand, since M is compact a minimizing measure must exist; therefore from the density of horocycle flow orbits we can easily deduce that there is a unique minimizing measure.

Appendix A

A.1 Absolutely continuous function

Recall that a function $f : [a, b] \to \mathbb{R}$ is *absolutely continuous* if for all $\epsilon > 0$ there is $\delta > 0$ such that

$$\sum_{i=1}^{n} |t_i - s_i| < \delta \quad \Longrightarrow \quad \sum_{i=1}^{n} |f(t_i) - f(s_i)| < \epsilon$$

whenever $]s_1, t_1[$,..., $]s_n, t_n[$ are disjoint intervals in $[a, b]$.

Proposition A.1.1. *The function $f : [a, b] \longrightarrow \mathbb{R}$ is absolutely continuous if and only if*

1. *The derivative $f'(t)$ exists for almost every $t \in [a, b]$;*
2. *$f' \in \mathcal{L}^1([a, b])$;*
3. *f satisfies the identity $f(t) = f(a) + \int_a^t f'(s)ds$;*

Proof. Suppose that f is absolutely continuous and define

$$\mu([s, t]) := f(t) - f(s) ;$$

we claim that μ defines a finite signed Borel measure on $[a, b]$. Indeed let $\mathcal{A}$ be the algebra of finite unions of intervals; the function μ can be extended to a σ-addictive function on $\mathcal{A}$. Moreover if B is a Borel set and $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ is a sequence such that $A_n \downarrow B$ then

$$\mu(B) := \lim_{n\to+\infty} \mu(A_n)$$

exists because $\mu(A_n \setminus A_m) \to 0$ when $n, m \to +\infty$. Observe that the properties of external measure and the absolute continuity of f imply that $\mu \ll \lambda$, where λ is the Lebesgue measure; now let $g = \frac{d\mu}{d\lambda}$ be the Radon-Nicodym derivative. Then $g \in \mathcal{L}^1$ and

$$f(t) - f(a) = \mu([a, t]) = \int_a^t g(s)ds ;$$

hence by the Lebesgue differentiation theorem

$$\lim_{h\to 0} \frac{f(t + h) - f(t)}{h} = \lim_{h\to 0} \frac{1}{h} \int_t^{t+h} g(s)\, ds = g(t) \quad \text{for a.e. } t \in [a, b] .$$

Conversely suppose that the function f satisfies the properties $1 - 3$ and let $\mu(A) := \int_A f' d\lambda$ (it is well defined because of 2); then by 3 we have

$$\mu([s, t]) = f(t) - f(s) \quad \forall\, s, t \in [a, b]$$

and hence $\mu \ll \lambda$, which implies that f is absolutely continuous. $\square$

Corollary A.1.2. *The function $f : [a, b] \longrightarrow \mathbb{R}$ is absolutely continuous if and only if there exists a function $g \in \mathcal{L}^1([a, b])$ such that*

$$f(t) = f(a) + \int_a^t g(s)\, ds .$$

A.2 The Fenchel and Legendre transforms

Recall that a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in \mathbb{R}^n$ and for all $\lambda \in [0,1]$ or, equivalently, if the set $\{(x,r) \in \mathbb{R}^n \times \mathbb{R} \mid r \geq f(x)\}$ is convex. Given a convex function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ the *subdifferential* of f at $x_0 \in \mathbb{R}^n$ is the set

$$\partial f(x_0) := \left\{p \in (\mathbb{R}^n)^* \;\middle|\; f(x) \geq p(x - x_0) + f(x_0)\right\};$$

its elements are called *subderivatives* (or *subgradients*) of f at x_0 and the hyperplanes

$$\{(x,r) \in \mathbb{R}^n \times \mathbb{R} \mid r = p(x - x_0) + f(x_0)\}, \quad p \in \partial f(x_0)$$

are called *supporting hyperplanes* for f at x_0, while the functional $p \in (\mathbb{R}^n)^*$ is called the *slope* of the hyperplane. Let us recall the following (see [4] for the proof)

Proposition A.2.1. *The following statements hold:*

1. *$\partial f(x) \neq \emptyset$ for every $x \in dom(f)$;*

2. *A finite convex function is continuous and Lebesgue almost everywhere differentiable;*

3. *If $\partial f(x) = \{p\}$ then f is differentiable at x and $f'(x) = p$;*

Definition A.2.2. *Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a convex function, then the **Fenchel transform** (or the convex dual) of f is the function $f^* : (\mathbb{R}^n)^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined by*

$$f^*(p) := \max_{x \in \mathbb{R}^n} \Big[p \cdot x - f(x)\Big]. \tag{A.1}$$

Observe that the function f admits a supporting hyperplane with slope $p \in (\mathbb{R}^n)^*$ if and only if $f^*(p) \neq +\infty$, indeed if $p \in \partial f(x_0)$ for some $x_0 \in \mathbb{R}^n$ then

$$f^*(p) = \max_{x \in \mathbb{R}^n} \Big[p \cdot x - f(x)\Big] \;\leq\; \max_{x \in \mathbb{R}^n} \Big[p \cdot x - p \cdot (x - x_0) - f(x_0)\Big] \;=\; p \cdot x_0 - f(x_0) \;<\; +\infty$$

where we have used the fact that if $p \in \partial f(x_0)$ then the inequality

$$f(x) \geq p \cdot (x - x_0) + f(x_0)$$

is satisfied by any $x \in \mathbb{R}^n$; conversely if $f^*(p) < +\infty$ then $f^*(p) = p \cdot x_0 - f(x_0)$ for some $x_0 \in \mathbb{R}^n$ and hence f admits a supporting hyperplane with slope p at x_0. Furthermore observe that if f is superlinear then f^* is finite on all $\mathbb{R}^n$, indeed in this case for each $p \in (\mathbb{R}^n)^*$ we have that the maximum in (A.1) is attained on a compact subset of $\mathbb{R}^n$; therefore if f is superlinear then f admits a supporting hyperplane with slope p for each $p \in (\mathbb{R}^n)^*$.

Proposition A.2.3. *The following hold:*

1. *The function f^* is convex;*

2. *If f and f^* are superlinear then $f^{**} = f$.*

3. *f is superlinear if and only if f^* is bounded on balls, i.e.*

$$f(x) \geq a|x| - b(a) \quad \forall x \in \mathbb{R}^n \quad \Longleftrightarrow \quad f^*(p) \leq b(|p|) \quad \forall p \in (\mathbb{R}^n)^*.$$

4. *If f is superlinear then the maximum in (A.1) is attained at some point $x \in \mathbb{R}^n$.*

Proof. Let us start proving statement 1; given $0 \leq \lambda \leq 1$ and $p_1, p_2 \in (\mathbb{R}^n)^*$ we have that

$$\begin{aligned} f^*(\lambda p_1 + (1-\lambda)p_2) &= \max_{x \in \mathbb{R}^n} \left[\big(\lambda p_1 + (1-\lambda)p_2\big) \cdot x - f(x) \right] \leq \\ &\leq \lambda \max_{x \in \mathbb{R}^n} \left[p_1 \cdot x - f(x) \right] + (1-\lambda) \max_{x \in \mathbb{R}^n} \left[p_2 \cdot x - f(x) \right] = \\ &= \lambda f^*(p_1) + (1-\lambda) f^*(p_2) \,. \end{aligned}$$

In order to prove item 2 observe that from (A.1) it follows that

$$f(x) \;\geq\; p \cdot x - f^*(p)$$

for all $x \in \mathbb{R}^n$ and for all $p \in (\mathbb{R}^n)^*$; hence

$$f(x) \;\geq \sup_{p \in (\mathbb{R}^n)^*} \left[p \cdot x - f^*(p) \right] = f^{**}(x) \,. \tag{A.2}$$

Conversely fix $x \in \mathbb{R}^n$ and consider $p_x \in \partial f(x)$; observe that such an element p_x effectively exists because by proposition A.2.1 the set $\partial f(x)$ is non empty. Since by the definition of subdifferential and subderivative we have that $f(y) \geq f(x) + p_x \cdot (y - x)$ for all $y \in \mathbb{R}^n$ we get

$$f^*(p_x) = \max_{y \in \mathbb{R}^n} \left[p_x \cdot y - f(y) \right] = p_x \cdot x - f(x)$$

and hence

$$f(x) = p_x \cdot x - f^*(p_x) \leq \max_{p \in (\mathbb{R}^n)^*} \left[p \cdot x - f^*(p) \right] = f^{**}(x) \,. \tag{A.3}$$

Combining equation (A.2) with equation (A.3) we get item 2. Now let us prove 3; by the superlinearity, for a suitable $b(|p|) \in \mathbb{R}$ we have that

$$f^*(p) = \max_{x \in \mathbb{R}^n} \left[p \cdot x - f(x) \right] \leq \max_{x \in \mathbb{R}^n} \left[p \cdot x - |p| \cdot x \right] + b(|p|) = b(|p|) \,.$$

Conversely if we suppose that $f^*(p) \leq b(|p|)$ for all $p \in (\mathbb{R}^n)^*$ then, given $a \in \mathbb{R}$ and $x \in \mathbb{R}^n$, there exists $p_x \in (\mathbb{R}^n)^*$ such that $|p_x| = a$ and $p_x \cdot x = |p_x||x| = a|x|$; hence

$$f(x) \geq f^{**}(x) = \max_{p \in (\mathbb{R}^n)^*} \left[p \cdot x - f^*(p) \right] \geq p_x \cdot x - b(|p_x|) = a|x| - b(a) \,.$$

In particular we have that if f is superlinear then f^* is finite; now let $p \in (\mathbb{R}^n)^*$. If $b > 0$ is such that $f(x) > (|p| + 1)|x| - b$ then

$$p \cdot x - f(x) < b - |x| < f^*(p) - 1 \qquad \text{for } |x| > b + 1 - f^*(p) \,,$$

hence the maximum in the definition (A.1) of the Fenchel transform is attained at some interior point x_p in the closed ball $\{|x| \leq b + 1 - f^*(p)\}$ and this proves statement 4. □

Therefore if $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex and superlinear then so is f^* and $f^{**} = f$; moreover we get

$$f^*(0) = -\min_{x \in \mathbb{R}^n} f(x) \,, \qquad f(0) = -\min_{p \in (\mathbb{R}^n)^*} f^*(p) \,.$$

In the convex and superlinear case one can also define the *Legendre transform* of f as

$$\mathcal{L} : \mathbb{R}^n \longrightarrow 2^{(\mathbb{R}^n)^*}, \qquad \mathcal{L}(x) := \{ p \in (\mathbb{R}^n)^* \mid p \cdot x = f(x) + f^*(p) \} = \partial f(x) \,.$$

Proposition A.2.4. *If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is C^2 and there exists $a > 0$ such that*

$$y \cdot f''(x) \cdot y \geq a|y|^2 \qquad \forall x, y \in \mathbb{R}^n$$

then the Legendre transform $\mathcal{L}$ is a C^1-diffeomorphism given by $\mathcal{L}(x) = d_x f$.

Proof. The function f is convex and superlinear because

$$f(x) = f(0) + \int_0^1 f'(sx)ds = f(0) + \int_0^1 \int_0^1 sx \cdot f''(tsx) \cdot x\, dtds \geq f(0) + \frac{1}{2}a|x|^2\,;$$

then by proposition A.2.3 f^* is superlinear and

$$\mathcal{L}(x) = \{p \in (\mathbb{R}^n)^* \mid f(x) = px - f^*(p)\} \neq \emptyset\,.$$

Moreover if $p \in \mathcal{L}(x)$ then by the definition of the Legendre transform and by (A.1) we get

$$p \cdot y - f(y) \leq p \cdot x - f(x)$$

for all $y \in \mathbb{R}^n$; now f is C^2 and this implies that $p = df(x) = \mathcal{L}(x)$ and hence $\mathcal{L}$ is differentiable and single valued. Moreover since $d\mathcal{L}(x) = f''(x)$ is non singular we get that $\mathcal{L}$ is a local diffeomorphism; at the same time the inequality

$$(y - x) \cdot [df(y) - df(x)] \;=\; \int_0^1 f''(sx + (1-s)y)\, ds \;>\; 0$$

implies that $\mathcal{L}$ is injective. Since the surjectivity of $\mathcal{L}$ follows immediately from item 4 in the proposition above this completes the proof. □

A.3 Birkhoff's ergodic theorem

Definition A.3.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space, then a **measure-preserving transformation** is a map $T : X \longrightarrow X$ such that*

1. *T is measurable, i.e. for all $A \in \mathcal{A}$ the set $T^{-1}(A) := \{x \in X \mid Tx \in A\}$ belongs to $\mathcal{A}$.*
2. *T preserves the measure μ, i.e. $\mu(T^{-1}(A)) = \mu(A)$ for each $A \in \mathcal{A}$.*

Suppose that T is a measure-preserving transformation of the probability space $(X, \mathcal{A}, \mu)$; given an element $B \in \mathcal{A}$ it is natural to wonder how often do orbits of a given point $x \in X$ visit the set B. If one pictures a point $x \in X$ being showed about by T which preserves μ it could be guessed that the frequency of visits to a set $B \in \mathcal{A}$ is perhaps $\mu(B)$; unfortunately this is false, but in "essence" it is true!! The problem which arises is that the space X could consist of two parts, say X_1 and X_2, and T could map points of X_1 only into X_1 and points of X_2 only into X_2; more drastically if T is the identity map (a very uninteresting case) then any given measure μ is preserved but any point inside B always remains in B whereas points outside B never get inside.

Theorem A.3.2 (Basic ergodic theorem)**.** *Let T be a measure-preserving transformation of the probability space $(X, \mathcal{A}, \mu)$ and let $B \in \mathcal{A}$; if*

$$S_n(x) := \;\{i \mid 0 \leq i < n\,,\; T^i x \in B\}$$

is the number of visits of the orbit of x to the set B up to time n and $A_n(x) := \frac{1}{n}S_n(x)$ is the average of these visits (up to time n) then the limit

$$A(x) := \lim_{n\to+\infty} A_n(x)$$

exists for μ-almost every $x \in X$.

Before proving the theorem we shall observe some important facts; clearly $A(x)$ is the average number of visits of the orbit of x to the set B and the Basic ergodic theorem implies that this average is well defined for a subset of X with full μ-measure. Moreover $A(\cdot)$ is invariant on X under T, in the sense that if $A(x)$ exists then so does $A(Tx)$ and $A(Tx) = A(x)$. We say that T is *ergodic* if any T-invariant measurable function $f : X \longrightarrow \mathbb{R}$ (i.e. f is measurable and satisfies $f(Tx) = f(x)$ for μ-almost every $x \in X$) is μ-almost everywhere constant. As a corollary of the proof of the Basic ergodic theorem we get that

$$\int_X A(x)\, d\mu(x) \;=\; \mu(B) \tag{A.4}$$

and hence if T is ergodic then it follows immediately that

$$A(x) \;=\; \mu(B) \tag{A.5}$$

for μ-almost every $x \in X$. Now let us prove the theorem; given $x \in X$ define

$$\overline{A}(x) := \limsup_{n\to+\infty} A_n(x)\,, \qquad \underline{A}(x) := \liminf_{n\to+\infty} A_n(x)\,.$$

Observe that $0 \le \underline{A}(x) \le \overline{A}(x) \le 1$ for each $x \in X$; therefore the theorem follows if

$$\int_X \overline{A}(x)\, d\mu(x) \;\le\; \mu(B)\,. \tag{A.6}$$

Indeed an analogous argument (or the same applied to $X \setminus B$ instead of B) shows that

$$\mu(B) \;\le\; \int_X \underline{A}(x)\, d\mu(x)$$

and hence $\int_X \left[\underline{A}(x) - \overline{A}(x)\right] d\mu(x) \ge 0$, together with $\underline{A} \le \overline{A}$, implies that $\overline{A}(x) = \underline{A}(x)$ for μ-almost every $x \in X$. Furthermore observe that from (A.6) we get equation (A.4) and hence equation (A.5) if T is ergodic. Now we prove the claim; fixed $\epsilon > 0$ consider

$$\tau(x) := \min\,\{n > 0 \mid A_n(x) \ge \overline{A}(x) - \epsilon\}\,;$$

if $\tau \in L^\infty(X)$ then there is $M \in \mathbb{N}$ such that $\tau(x) \le M$ for μ-a.e. $x \in X$ and hence the inequality

$$S_n(x) \;\ge\; (n - M)(\overline{A}(x) - \epsilon)$$

follows by using τ to decompose the orbit of x up to time n into pieces on each of which the average number of visits of B is at least $\overline{A}(x) - \epsilon$ and the piece left over has length bounded by M. Therefore dividing by n, using the fact that T is μ-preserving and integrating over X we get

$$\mu(B) \;=\; \int_X A_n(x)\, d\mu(x) \;\ge \int_X \left(1 - \frac{M}{n}\right)(\overline{A}(x) - \epsilon)\, d\mu(x) \;\overset{n\to+\infty}{\longrightarrow} \int_X \overline{A}(x)\, d\mu(x) - \epsilon\,.$$

In general $\tau \notin L^\infty(X)$ but however the proof in the general case is more or less the same as that above and consists into apply the same argument to a slight modification of the set B. Suppose $\epsilon > 0$ fixed and let $M \in \mathbb{N}$ such that

$$\mu(\{x \in X \mid \tau(x) > M\}) \;<\; \epsilon\,;$$

observe that such an M effectively exists because otherwise we would get a contradiction with the definition of $\overline{A}(\cdot)$. Now define the set

$$B' := \; B \cup \{x \in X \mid \tau(x) > M\}$$

and denote by $S'_n(x)$ the number of visits to B' of the orbit of x; as before one can prove that $S'_n(x) \ge (n - M)(\overline{A}(x) - \epsilon)$ and hence we finally get

$$\mu(B) \;\ge\; \mu(B') - \epsilon \;\ge \int_X \overline{A}(x)\, d\mu(x) - 2\epsilon\,.$$

The Basic ergodic theorem answers to the question "how often do orbits of a given point $x \in X$ visit the set B", but there are lot of questions left open by this theorem; for instance the work of Boltzmann and Gibbs on statistical mechanics raised the following mathematical problem: given a measure-preserving map T of a probability space $(X, \mathcal{A}, \mu)$ and an integrable function $f : X \longrightarrow \mathbb{R}$, find conditions under which the limit

$$\lim_{n\to+\infty} \frac{f(x) + f(Tx) + \ldots + f(T^{n-1}x)}{n} \tag{A.7}$$

exists and it is constant almost everywhere. Observe that the Basic ergodic theorem solves the case $f = I_B$, where I_B denote the charateristic function of a measurable set $B \in \mathcal{A}$. In 1931 Birkhoff proved that for any T anf f as above the limit in (A.7) exists almost everywhere; from this result he showed that a necessary and sufficient condition for its value to be constant almost everywhere is that the transformation T is ergodic, although the definition of ergodicity is not exactly the same that we gave before (but however equivalent; see the end of this section). Birkhoff did not close the problem that motivated it, since for the maps that occur in statitical mechanics ergodicity could not then be proved; only in the sixties did the results of Sinai, and later those of Bunimovich, imply that maps analogous to the ones studied in statistical mechanics are ergodic .

Theorem A.3.3 (Birkhoff's ergodic theorem)**.** *Let $(X, \mathcal{A}, \mu)$ be a probability space and $T : X \longrightarrow X$ be a measure-preserving transformation, then the following hold:*

1. *If $f \in L^1(X)$ then the limit*
$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x)$$
exists for μ-almost every $x \in X$.

2. *If $f \in L^p(X)$, $1 \le p < +\infty$, then the function $\tilde{f}$ defined by*
$$\tilde{f}(x) := \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x)$$
is in $L^p(X)$ and it is T-invariant; moreover the partial sums converge to $\tilde{f}$ in L^p, i.e.
$$\Big\| \tilde{f} - \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j \Big\|_p \longrightarrow 0\,.$$

3. *For every $f \in L^p(X)$ we have $\displaystyle\int_X \tilde{f}\, d\mu = \int_X f\, d\mu\,.$*

Corollary A.3.4. *For every $A, B \in \mathcal{A}$ the limit*
$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap B)$$
exists.

We refer to [22] for the proof; as already pointed out there are other equivalent definitions of ergodicity, for example one can define T to be ergodic if every T-invariant set has measure 0 or 1. It is pointless to discuss here about these equivalent definitions; therefore we end this section just enouncing the following

Proposition A.3.5. *The following are equivalent:*

1. *T is ergodic, i.e. every T-invariant set has measure 0 or 1.*
2. *If $f \in \mathcal{L}^1(X)$ is T-invariant then f is constant μ-almost everywhere.*
3. *If $f \in \mathcal{L}^p(X)$ is T-invariant then f is constant μ-almost everywhere.*
4. *For every $f \in \mathcal{L}^1(X)$ we have $\tilde{f} = \int_X f\, d\mu$ almost everywhere.*
5. *For every $A, B \in \mathcal{A}$ we have*
$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) \;=\; \mu(A) \cdot \mu(B)\,.$$

Appendix B

B.1 Fundamental domains

Definition B.1.1. *If X is a topological space and $G = Hom(X)$ is the set of homeomorphisms of X into itself then a subgroup $G_0 < G$ is said to act* ***discontinuously*** *if for every compact subset $K \subseteq X$ the intersection $G(K) \cap K$ is non empty only for finitely many $g \in G_0$.*

In particular if G_0 acts discontinuously on X then the stabilizer

$$\mathrm{Stab}_{G_0}(x) \;=\; \{g \in G_0 \mid g(x) = x\}$$

is finite for every $x \in X$. If Γ is a topological group then a subgroup $\Gamma_0 < \Gamma$ is *discrete* if every point in Γ_0 is isolated (in the relative topology on Γ_0); in particular if Γ_0 is discrete then each sequence $\{g_i\} \subseteq \Gamma_0$ that converges to $\gamma \in \Gamma$ is definitively constant, indeed $g_{n+1}g_n^{-1} \in \Gamma_0$ and converges to the (isolated) identity. In the particular case

$$\Gamma = SL(2,\mathbb{C}), \qquad G = \mathrm{Aut}\,(S^2)$$

there is a homomorphism of groups $h : \Gamma \longrightarrow G$ with kernel $\{\pm I\}$; in this case we have

Theorem B.1.2. *Let D be a disc (or half-plane) in S^2 and let $G_0 < G$ be a subgroup such that $g(D) = D$ for every $g \in G_0$; moreover let $\Gamma_0 < \Gamma$ which projects to G_0. Then Γ_0 is discrete if and only if G_0 acts discontinuously on D.*

A *Fuchsian group* is a group G_0 of Möbius transformations acting discontinuously in some disc $D \subseteq S^2$; we may always suppose, up to componing with suitable conformal automorphisms of S^2 into itself, that $D = \Delta$ or $D = \mathbb{H}$. If G_0 is a Fuchsian group then a *fundamental domain* F for G_0 is an open subset of Δ (or $\mathbb{H}$) such that

$$\bigcup_{g \in G_0} g(\overline{F}) \;=\; \Delta\,, \qquad g(F) \cap h(F) = \emptyset \quad \text{if } \; g \neq h$$

where $\overline{F}$ denotes the closure of F relative to Δ (or $\mathbb{H}$). Thus a subset $F \subseteq \Delta$ is a fundamental domain if every point of Δ lies in the closure of some image $g(F)$ and if two distinct images of F do not overlap; in this case we say that F *tesselate* Δ. For example the set

$$F = \{x + iy \mid x \in (0,1).y > 0\}$$

is a fundamental domain for $< z \longmapsto z + 1 >$ acting on $\mathbb{H}$. In general F will be a poligon (in the appropriate geometry) and there will be side-pairing maps mapping one side of the poligon onto another; however observe that:

- The subset $\{z \in \mathbb{H} \mid 1 < x < 2\}$ is a poligon in $\mathbb{H}$ with the side-pairing map $g : z \longmapsto 2z$ but it is not a fundamental domain for $< g >$.
- The subset $\{z \in \Delta \mid 0 < \arg{(z)} < \frac{4}{7}\pi\}$ is a poligon in Δ with side-pairing map $g : z \longmapsto e^{\frac{4}{7}\pi i}$ but it is not a fundamental domain for $< g >$.

B.2 The hyperbolic plane $\mathbb{H}$

The hyperbolic plane is the subset of the real plane $\mathbb{R}^2$ composed by the points with positive second coordinate and endowed with the riemannian metric

$$g(x,y) \;=\; \frac{1}{y^2}\,(dx\otimes dx+dy\otimes dy)\,.$$

Notice that the metric is conform to the Euclidean one. Of course $\mathbb{H}$ can also be seen as a subset of the complex plane $\mathbb{C}$; we will use indistinctly both of these characterizations. We are interested in to determine the isometries of $\mathbb{H}$ and in to compute its geodesics, but before we shall explain where this specific Riemannian metric comes from. Let Δ be the unit disc in $\mathbb{C}$ and denote by $\mathrm{Aut}\,(\Delta)$ the set of conformal (i.e. angles preserving) maps of Δ into itself; its elements are called *conformal automorphisms* (or simply automorphisms, the definition is analogous for $\mathbb{H}$ and for the Riemann sphere S^2). We will see introducing a suitable metric ρ on Δ, with respect to which all the conformal automorphisms of Δ are isometries, that the metric on Δ induces naturally the Riemannian metric g on $\mathbb{H}$ and that the geodesics of $(\mathbb{H},g)$ can be easily deduced from the ones of (Δ,ρ). Furthermore we will show that the conformal automorphisms of $\mathbb{H}$ are "almost" all the isometries of $\mathbb{H}$ (the same holds for Δ) in a sense that we are going to specify at the end of this section; as we will see later in this section, it is also possible to classify the automorphisms of the hyperbolic plane with respect to their behaviour. As first step we shall deduce the conformal automorphisms of Δ and $\mathbb{H}$ from the ones of the Riemann sphere; it is well known that the conformal automorphisms of S^2 (also called biolomorphisms) are given by the *Möbius transformations*

$$f(z) \;=\; \frac{az+b}{cz+d}\,,\quad a,b,c,d\in\mathbb{C}\,,\quad ad-bc\neq 0\,. \tag{B.1}$$

Any Möbius transformations is conformal because it is obtained by composing translations, inversions, rotations and homotheties that are angles preserving (although they are not in general length and area preserving). To each Möbius tranformation as in (B.1) we can associate the matrix

$$A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in GL(2,\mathbb{C})\,;$$

every $f\in\mathrm{Aut}\,(S^2)$ is a bijection and f^{-1} has associated matrix A^{-1}. Moreover if $f,g\in\mathrm{Aut}\,(S^2)$ have associated matrices A,B then a simple calculation shows that the composition $g\circ f$ has associated matrix BA; this implies that the map

$$h: GL(2,\mathbb{C})\longrightarrow \mathrm{Aut}\,(S^2)\,,\quad \begin{pmatrix} a & b\\ c & d\end{pmatrix}\longmapsto f(z)=\frac{az+b}{cz+d}\,,$$

is a homomorphism of groups. Its kernel is given by $\{\lambda I \mid \lambda\in\mathbb{C}\setminus\{0\}\}$; hence by the homomorphism theorem we get an isomorphism

$$PGL(2,\mathbb{C}) \;=\; GL(2,\mathbb{C})\big/(\lambda I)\;\cong\;\mathrm{Aut}\,(S^2)\;\cong\; SL(2,\mathbb{C})\big/(\pm I)\;=\;PSL(2,\mathbb{C})\,.$$

Now it is easy to see that an element of $\mathrm{Aut}\,(S^2)$ preserves $\mathbb{H}$ if and only if all coefficients a,b,c,d are real and $ad-bc>0$; indeed the only elements that leave the real axis fixed are

$$f(z) \;=\; \frac{az+b}{cz+d}\qquad a,b,c,d\in\mathbb{R} \tag{B.2}$$

and hence it suffices to see where i is mapped by such an f. With a simple computation one can show that i is mapped by such an f into a point of $\mathbb{H}$ if and only if $ad-bc>0$; observe that if $ad-bc<0$ then f maps $\mathbb{H}$ into the lower half-plane. The conformal map

$$f(z) \;=\; \frac{z-i}{z+i}$$

is a bijection between $\mathbb{H}$ and Δ; therefore the set of Möbius transformations of Δ into itself is transitive because so is the set of Möbius transformations of $\mathbb{H}$ (we can map i into every point of $\mathbb{H}$ using an element as in (B.10) such that $ad-bc>0$). Now it is just an easy consequence of the

Schwartz' Lemma the fact that conformal maps from Δ into itself (and hence from $\mathbb{H}$ into itself) are only given by Möbius transformations. Observe that the homomorphism of group

$$SL(2,\mathbb{R}) \longrightarrow \mathrm{Aut}\,(\mathbb{H})\,, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto f(z) = \frac{az+b}{cz+d}$$

has kernel $\{\pm I\}$ and defines, by the homomorphism theorem, an isomorphism between $PSL(2,\mathbb{R})$ and $\mathrm{Aut}\,(\mathbb{H})$. Similarly one can show that each $g \in \mathrm{Aut}\,(\Delta)$ is of the form

$$g(z) \;=\; \frac{az+\bar{c}}{cz+\bar{a}}\,, \qquad |a|^2-|c|^2 = 1\,;$$

if we take the derivative of g with a simple computation we get the relation

$$\frac{1-|g(z)|^2}{1-|z|^2} = \frac{|cz+\bar{a}|^2-|az+\bar{c}|^2}{|cz+\bar{a}|^2(1-|z|^2)} = \frac{(|a|^2-|c|^2)(1-|z|^2)}{|cz+\bar{a}|^2(1-|z|^2)} = \frac{1}{|cz+\bar{a}|^2} \;=\; |g'(z)| \tag{B.3}$$

which holds for every conformal automorphism $g \in \mathrm{Aut}\,(\Delta)$. This suggests to define the function

$$\lambda : \Delta \longrightarrow (0,+\infty)\,, \qquad z \longmapsto \frac{2}{1-|z|^2}$$

in order to obtain the following identity

$$\lambda(g(z))|g'(z)| \;=\; \frac{2}{1-|g(z)|^2}\cdot\frac{1-|g(z)|^2}{1-|z|^2} \;=\; \lambda(z) \tag{B.4}$$

for each $g \in \mathrm{Aut}\,(\Delta)$. One can also define λ by just requirring that it satisfies (B.4); this point of view will be used later in this section when we will generalize these arguments to a generic open simply connected subset of $\mathbb{C}$. The *hyperbolic length* of a curve $\gamma : I \longrightarrow \Delta$ is defined by

$$L(\gamma) := \int_\gamma \lambda(z)|dz|\,;$$

observe that L is invariant under $\mathrm{Aut}\,(\Delta)$ because of (B.4), indeed

$$L(g\circ\gamma) \;=\; \int_{g\circ\gamma}\frac{2}{1-|w|^2}|dw| \;=\; \int_\gamma \frac{2|g'(z)|}{1-|g(z)|^2}|dz| \;=\; \int_\gamma\frac{2}{1-|z|^2}|dz| \;=\; L(\gamma)\,.$$

Theorem B.2.1. *The function $\rho : \Delta\times\Delta \longrightarrow \mathbb{R}$ defined by*

$$\rho(z,w) := \;\inf\;\{L(\gamma)\mid \gamma : I \longrightarrow \Delta\,,\; \gamma(0)=z\,,\; \gamma(T)=w\}\,, \qquad \forall\, z,w\in\Delta$$

is a metric on Δ and the conformal automorphisms are isometries of (Δ,ρ).

Proof. Each $g \in \mathrm{Aut}\,(\Delta)$ is an isometry for (Δ,ρ) because, as already pointed out, the hyperbolic length is invariant under $\mathrm{Aut}\,(\Delta)$; moreover all the properties of a metric, except

$$z \neq w \quad\Longrightarrow\quad \rho(z,w) > 0\,, \tag{B.5}$$

follows obviously from the definition of ρ. In order to prove (B.5) let us first compute $\rho(0,z)$ for a generic $z \in \Delta$; more precisely we want to show that

$$\rho(0,z) \;=\; \log\left(\frac{1+|z|}{1-|z|}\right)\,. \tag{B.6}$$

If g is a rotation about the origin that maps z onto $|z|$ then we have $\rho(0,z) = \rho(0,|z|)$; hence we can suppose $z = x \in (-1,1)$. Let now $\gamma(t) = \alpha(t)+i\beta(t)$ be a curve on Δ joining 0 to x, then

$$L(\gamma) \;=\; \int_0^1 \frac{2\sqrt{\dot\alpha^2+\dot\beta^2}}{1-(\alpha^2+\beta^2)}\,dt \;\geq \int_0^1\frac{2\sqrt{\dot\alpha^2}}{1-\alpha^2}\,dt \;\geq \int_0^1\frac{2\dot\alpha}{1-\alpha^2}\,dt \;=\int_0^x \frac{2}{1-s^2}\,ds \;=\; \log\left(\frac{1+x}{1-x}\right)$$

and equality holds when $\dot\alpha > 0$ and $\beta = 0$. This computation also proves that

$$\rho(0,z) > 0 \quad \text{if} \quad z \neq 0\,, \qquad \rho(0,z) \longrightarrow +\infty \quad \text{if} \quad |z| \longrightarrow 1\,.$$

Now we want an analogous expression for $\rho(z,w)$; clearly it suffices to apply a conformal automorphism g that maps w onto 0 and then use (B.6). Observe that such an automorphism g exists because Aut (Δ) is transitive over Δ and in particular is given by

$$g(\zeta) \;=\; \frac{\zeta - w}{1-\zeta\bar w}\,.$$

Thus

$$\rho(z,w) \;=\; \rho(0,g(z)) \;=\; \log\left(\frac{1+|g(z)|}{1-|g(z)|}\right) \;=\; \log\left(\frac{|1-z\bar w| + |z-w|}{|1-z\bar w| - |z-w|}\right)\,;$$

hence in particular

$$\cosh\rho(z,w) \;=\; \frac{|1-z\bar w|^2 + |z-w|^2}{|1-z\bar w|^2 - |z-w|^2}\,. \tag{B.7}$$

and this completes the proof because (B.5) follows obviously from (B.7). □

The metric ρ defined above is called the *hyperbolic metric*; given a positive real number $r > 0$ we can define the *hyperbolic circle* of (hyperbolic) radius r and (hyperbolic) center 0 as

$$C := \{z \in \Delta \mid \rho(z,0) = r\}\,;$$

by equation (B.6) we have that C is equal to the Euclidean circle $C_R = \{z \in \Delta \mid |z| = R\}$ where

$$r = \rho(0,R) = \log\left(\frac{1+R}{1-R}\right)\,.$$

Analogously we can define the hyperbolic circle $C(w,r)$ of (hyperbolic) radius r and (hyperbolic) center $w \in \Delta$; using a conformal automorphism that maps w onto 0 we get tha any iperbolic circle is Euclidean with radius R as above, although in general its Euclidean center is not w (this is however true in the case $w = 0$). The argument above also clearly applies to the case of any *hyperbolic disc*

$$D(w,r) = \{z \in \Delta \mid \rho(z,w) < r\}$$

by first observing that $D(0,r)$ is the Euclidean disc of radius R centered at the origin where R is as above and then extending to a generic w with a conformal automorphism that maps w into the origin. Therefore we have proved the following

Proposition B.2.2. *The following hold:*

1. *The class of hyperbolic circles in Δ coincides with the class of Euclidean circles in Δ.*

2. *The class of hyperbolic discs in Δ coincides with the class of Euclidean discs in Δ.*

3. *The hyperbolic topology on Δ coincides with the Euclidean one.*

Observe that item 3 follows obviously from the fact that hyperbolic discs coincide with Euclidean ones. The length of the hyperbolic circle $C(w,r)$ is the same as that of the Euclidean circle C_R with R as above, hence

$$L(C(w,r)) \;=\; L(C_R) \;=\int_0^{2\pi} \frac{2R}{1-R^2}\,d\vartheta \;=\; \frac{4\pi R}{1-R^2} \;=\; 2\pi\sinh r \tag{B.8}$$

In analogy to the Euclidean case one can also define the *hyperbolic area* of a subset $E \subseteq \Delta$ as

$$A(E) := \int_E \lambda(z)^2\,dxdy\,;$$

as it was for the hyperbolic length we have that the (hyperbolic) area is invariant under $\mathrm{Aut}\,(\Delta)$ and a simple calculation shows that the area of $D_r = \{z \in \Delta \mid \rho(0,z) < r\}$ is

$$A(D_r) \;=\; 4\pi \sinh^2\left(\frac{r}{2}\right). \tag{B.9}$$

Observe that in the hyperbolic case we have

$$L(C_r) \;\sim\; A(D_r) \qquad \text{as} \qquad r \longrightarrow +\infty$$

in contrast with the Euclidean case. Now consider two distinct points $z, w \in \Delta$; we say that a curve $\gamma : I \longrightarrow \Delta$ joining z to w is a *geodesic* if it satisfies

$$L(\gamma) \;=\; \rho(z,w)\,;$$

in other words γ is a geodesic joining z to w if it minimizes the hyperbolic length between z and w. The correct definition should be that γ is a geodesic if it *locally* minimizes the length; in general this two definitions do not coincide (consider for instance the sphere S^n) but in this case they are equivalent. For the general definition of geodesics in Riemannian geometry see any book of differential geometry, as for instance [3]. We do not know a priori if geodesics actually exist or not, but it is not difficult to prove that the answer to this question is affirmative. Indeed we pointed out in the proof of theorem B.2.1 that if $\gamma(t) = \alpha(t) + i\beta(t)$ is a curve joining 0 to $x \in (-1,1)$ then

$$L(\gamma) \;\geq\; \log\left(\frac{1+x}{1-x}\right)$$

where the equality holds if and only if $\beta \equiv 0$ and $\dot{\alpha} > 0$. Therefore in the case $z = 0$ and $w = x$ we have that a geodesic γ is given by the segment joining 0 to x (contained in the real axis); on the other hand it is clear that any other geodesic is just a reparametrization of γ, i.e. the geodesic is unique up to reparametrizations. Hereafter all the geodesics are supposed to be parametrized by arc-length; observe that the real axis is the unique Euclidean circle (lines are thought as circles with infinit radius) that is orthogonal to $\partial\Delta$ in $(-1,0)$ and in $(1,0)$. Since all of these properties observed are preserved by conformal automorphisms we get that there is a unique geodesic joining two distinct points $z, w \in \Delta$ and this geodesic is given by the simple arc (contained in Δ) of the Euclidean circle through z and w that is orthogonal to $\partial\Delta$. We have just proved the following

Theorem B.2.3. *Let $z, w \in \Delta$ be two distinct points and let C be the unique Euclidean circle through z and w that is orthogonal to $\partial\Delta$; then for all $\gamma : I \longrightarrow \Delta$ joining z and w we have*

$$L(\gamma) \;\geq\; \rho(z,w)\,.$$

Moreover the equality holds if and only if γ is the simple arc of C joining z and w in Δ; in other words the geodesics of the hyperbolic disc (Δ, ρ) are the (Euclidean) segments through the origin and the simple arcs, contained in Δ, of the circles orthogonal to $\partial\Delta$.

A geodesic of Δ is uniquely determined by its intersections ξ, η with the boundary $\partial\Delta$. Hence the set of geodesics can be seen as

$$\{(\xi,\eta) \in \partial\Delta \times \partial\Delta \mid \xi \neq \eta\}$$

where a geodesic is uniquely determined by (ξ, η). Another way to determine a geodesic is to give a point $w \in \Delta$ and a unit vector $e^{i\vartheta} \in \partial\Delta$; in this case it is possible that two different pair $(w, e^{i\vartheta})$, $(\xi, e^{i\eta})$ determine the same geodesic. However, although the representation is not unique, the set of geodesics can also be seen as the set of pairs

$$\{(w, e^{i\vartheta}) \in \Delta \times \partial\Delta\}\,.$$

Now we want to generalize all that we have done so far to any open and simply connected set $D \subseteq \mathbb{C}$, $D \neq \mathbb{C}$. Thus let D an open and simply connected subset of $\mathbb{C}$ and let $f : \Delta \longrightarrow D$ be a conformal map; this map exists because by *Riemann's theorem* every open and simply connected subset of the complex plane that is different from $\mathbb{C}$ is biolomorphic to Δ (see for instance [16]). As done before we can define the function $\lambda_D : D \longrightarrow (0, +\infty)$ by requirring that it satisfies the property

$$\lambda_D(f(z)) \cdot |f'(z)| \; = \; \lambda(z) \qquad \forall\, z \in \Delta \tag{B.10}$$

where $\lambda(\cdot)$ is defined by (B.4) and the metric ρ_D as done in theorem B.2.1 by changing Δ with D. Hence f is an isometry from (Δ, ρ) to (D, ρ_D); this obviously applies also to the special case of the *hyperbolic plane*. Recall that the conformal map

$$f^{-1}(z) := \frac{z-1}{z+i}$$

maps $\mathbb{H}$ into Δ and the boundary $\partial\mathbb{H} = \mathbb{R} \cup \{\infty\}$ into $\partial\Delta$.

Lemma B.2.4. *We have that* $\lambda_{\mathbb{H}}(z) = \frac{1}{y}$ *, where* $z = x + iy \in \mathbb{H}$.

Proof. It suffices to apply equation (B.10) to $z = f^{-1}(w)$, $w = x + iy \in \mathbb{H}$; recalling the formula of the derivative of the inverse function we get

$$\begin{aligned} \lambda_{\mathbb{H}}(w) \;&=\; \frac{\lambda(f^{-1}(w))}{|f'(f^{-1}(w))|} \;=\; \lambda(f^{-1}(w)) \cdot |(f^{-1})'(w)| \;= \\ &=\; \lambda\left(\frac{w-1}{w+1}\right) \cdot \left|\frac{-2i}{(w+i)^2}\right| \;=\; \frac{2}{|w+i|^2} \cdot \frac{2\,|w+i|^2}{|w+i|^2 - |w-i|^2} \;= \\ &=\; \frac{4}{|w+i|^2 - |w-i|^2} \;=\; \frac{4}{|x+y(1+i)|^2 - |x+y(1-i)|^2} \;=\; \frac{1}{y}\,. \end{aligned}$$

□

Corollary B.2.5. *The hyperbolic plane* $\mathbb{H}$ *endowed with the riemannian metric*

$$g \;=\; \frac{1}{y^2}\,(dx \otimes dx + dy \otimes dy)$$

is isometrically equivalent to the hyperbolic disc; moreover since the geodesics of Δ *are mapped into those of* $\mathbb{H}$ *by the conformal map* f *above, we get that the geodesics of* $\mathbb{H}$ *are the half-circles (and vertical lines) orthogonal to the* x*-axis. The vertical geodesics of* $\mathbb{H}$ *are parametrized by*

$$\gamma_x(t) = (x, e^t)\,;$$

therefore each geodesic of $\mathbb{H}$ *has all* $\mathbb{R}$ *as its domain of definition and this fact, by the Hopf-Rinow theorem, is equivalent to say that the metric on* $\mathbb{H}$ *(and hence also the metric on* Δ*) is complete.*

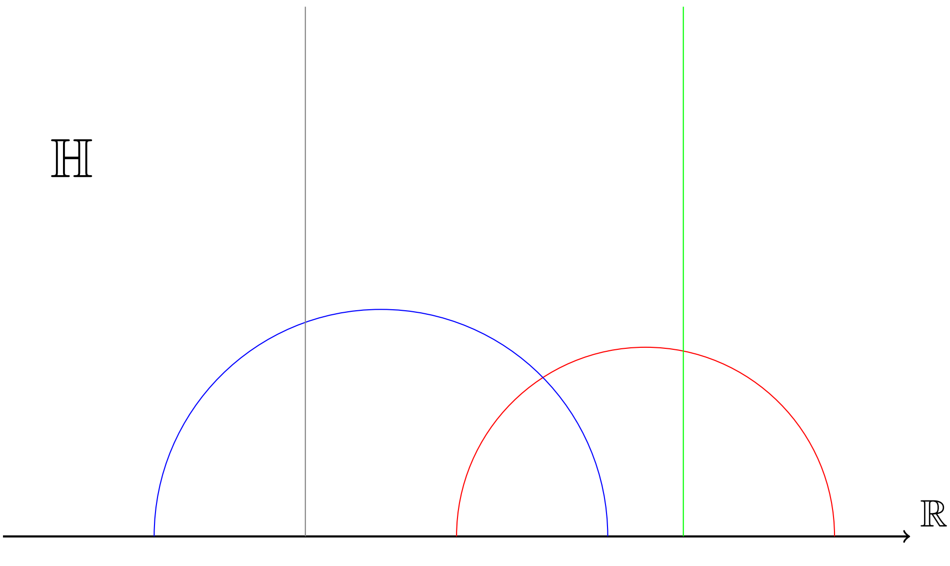


Figure B.1 *The geodesics of the hyperbolic plane are vertical lines or half-circles hortogonal to* $\mathbb{R}$.

B.3 Isometries of the hyperbolic plane

We already know that automorphisms of $\mathbb{H}$ (and of Δ) are isometries; thus the following question arises naturally: these are the only isometries or there is something else? The answer turn to be "almost" affirmative in a sense that we are going to specify now.

Theorem B.3.1. *Every isometry of* (Δ, ρ) *is given by* $z \longmapsto g(z)$ *or* $z \longmapsto g(\bar{z})$ *with* $g \in Aut(\Delta)$.

Proof. The proof is very simple and consists, given an isometry h, into componing h with suitable automorphisms of Δ in order to get the identity or the $z \longrightarrow \bar{z}$ function. Since Aut (Δ) is transitive over Δ, given an isometry h, there is an automorphism g_1 such that $g_1(h(0)) = 0$, i.e. such that the composition $h_1 = g_1 \circ h$ fixes the origin. Now since h_1 is still an isometry we get

$$\rho(\,0, \frac{1}{2}\,) \;=\; \rho(\,0, h_1(\frac{1}{2})\,) \qquad \Longrightarrow \qquad \left| h_1(\frac{1}{2}) \right| \;=\; \frac{1}{2}$$

and hence we can find a rotation g_2 about the origin such that $g_2(h_1(\frac{1}{2})) = \frac{1}{2}$, i.e. such that $h_2 = g_2 \circ h_1$ fixes the origin and the point $\frac{1}{2}$. Now it is easy to see that h_2 fixes all the points in the real segment $(-1, 1)$ and this implies that either $h_2(z) = z$ or $h_2(z) = \bar{z}$. □

The same argument also applies to $\mathbb{H}$; given an isometry h we can suppose that h fixes the points i and $2i$. Indeed Aut $(\mathbb{H})$ is transitive over $\mathbb{H}$ and hence there is an automorphism g_1 such that $h_1 = g_1 \circ h$ fixes i; moreover if $f : \mathbb{H} \longrightarrow \Delta$ is the standard isometry then f maps i into the origin and $2i$ into the point $\frac{1}{3}$. Therefore, since h_1 is still an isometry of $\mathbb{H}$, we get

$$\rho(0, f(h_1(2i))) \;=\; \rho_{\mathbb{H}}(i, h_1(2i)) \;=\; \rho_{\mathbb{H}}(i, 2i) \;=\; \rho(0, f(2i))$$

and hence up to componing with a rotation about the origin in Δ we may suppose that $f(h_1(2i)) = \frac{1}{3}$, that is $h_1(2i) = 2i$. Since h_1 fixes i and $2i$ we have that h_1 fixes each point in the upper part of the imaginary axis $i\mathbb{R}^+$ and this implies that each isometry h of $\mathbb{H}$ is either $z \longmapsto g(z)$ or $z \longmapsto g(\bar{z})$ for a suitable $g \in$ Aut $(\mathbb{H})$, and we are done. The following theorem is about the fixed points of an automorphism of the hyperbolic plane.

Theorem B.3.2. *Let* $h \neq id$ *be an automorphism of* $\mathbb{H}$*, then exactly one of the following holds:*

1. *h has a singled fixed point and this lies on $\partial\mathbb{H}$.*
2. *h has two fixed points and these both lie on $\partial\mathbb{H}$.*
3. *h has a single fixed point on $\mathbb{H}$ and none on $\partial\mathbb{H}$.*

Proof. The point ∞ is fixed if and only if $c = 0$ and in this case the equation

$$z \;=\; g(z) \;=\; \frac{az + b}{d}$$

has at most one (finite) real solution; so, since the coefficient are real, the equation $g(z) = z$ has either two (maybe coincident) real solutions or two complex conjugated solutions. □

Of course an analogous theorem holds for the hyperbolic unit disc (Δ, ρ). Now we want to classify automorphisms of the Riemann sphere (and of $\mathbb{H}$ and Δ) with respect to their behaviour; recall that Aut (S^2) is canonically isomorphic to $PSL(2, \mathbb{C})$; thus is well defined the application

$$\tau : \mathrm{Aut}\,(S^2) \longrightarrow \mathbb{C}\,, \qquad \tau(g) \;=\; \mathrm{tr}\,(A)^2 \;=\; \mathrm{tr}\,(-A)^2$$

where A is the matrix associated to g and tr (A) is the trace of A. It follows from elementary facts of linear algebra that the application τ is invariant over conjugation, that is $\tau(ghg^{-1}) = \tau(h)$ and conversely it is possible to prove that if two automorphisms have the same trace then they are conjugated. Therefore we have the following

Lemma B.3.3. *Two automorphisms* $f, g \in Aut(S^2)$ *are conjugated if and only if* $\tau(f) = \tau(g)$.

Now we introduce the following definitions:

- An automorphism $g \in \mathrm{Aut}\,(S^2)$ is *parabolic* if it is conjugated to $z \longmapsto z + b$, $b \neq 0$.
- An automorphism $g \in \mathrm{Aut}\,(S^2)$ is *elliptic* if it is conjugated to $z \longmapsto ze^{i\vartheta}$, $\vartheta \neq 0$.
- An automorphism $g \in \mathrm{Aut}\,(S^2)$ is *hyperbolic* if it is conjugated to $z \longmapsto kz$ with $k > 0$, $k \neq 1$.

Analogous definitions hold for $\mathbb{H}$ and Δ; we shall see that parabolic, iperbolic and elliptic automorphisms are three distinct classes (i.e. a parabolic automorphism can not be elliptic, ...) and that every automorphism of the Riemann sphere is of one of this three types; in other words $g \in \mathrm{Aut}\,(S^2)$ is either parabolic or elliptic or hyperbolic. We will get this result proving that each of those classes has a different range of possible τ-function values and that the union of their ranges is all $\mathbb{R}^+$; observe that translations are all conjugated to each other (because they all have the same trace) while rotations and homotheties do not. Indeed two rotations are conjugated if and only if they have the same order while two homotheties are conjugated if and only if their factor k, k' satisfies $k = (k')^{-1}$. Thus parabolic isometries are all in the same conjugated class while there are different conjugated classes of hyperbolic and elliptic isometries.

Theorem B.3.4. *Let $g \neq id$ be an automorphism of S^2 into itself, then the following are equivalent:*

1. *g is parabolic.*
2. *g has a unique fixed point in S^2.*
3. *$\tau(g) = 4$.*

Proof. If g is parabolic then it has exactly one fixed point $w \in S^2$ because a translation does; so let us suppose that g has a unique fixed point. We have to distinguish three cases:

- $\underline{w = \infty}$ In this case g is of the form
$$g(z) \;=\; \frac{az+b}{d}\,, \qquad A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in PSL(2,\mathbb{C})$$
with $ad = 1$; hence $a = 1$ (otherwise g would have two fixed points) and thus $d = 1$, $\tau(g) = 4$.
- $\underline{w \neq 0, \infty}$ In this case the coefficient b, c are non zero and $g(z) = z$ has a unique solution if and only if $cz^2 + (d-a)z - b = 0$ has a unique solution, that is if and only if
$$\begin{aligned} 0 &= (a-d)^2 + 4bc \;=\; a^2 - 2ad + d^2 + 4bc \;=\; a^2 - 2ad + d^2 + 4ad - 4 \;= \\ &= a^2 + 2ad + d^2 - 4 \;=\; (a+d)^2 - 4 = \tau(g) - 4\,. \end{aligned}$$
- $\underline{w = 0}$ In this case $b = 0$ and an analogous calculation as that above shows that $w = 0$ is the unique solution of $g(z) = z$ if and only if $c \neq 0$, $d = a$, i.e. if and only if $\tau(g) = 4$.

Now suppose $\tau(g) = 4$; since automorphisms are conjugated if and only if they have the same trace, g is conjugated to an automorphism f_A with associated matrix
$$A \;=\; \begin{pmatrix} a & b \\ 0 & \frac{1}{a} \end{pmatrix}$$
and the condition $\tau(g) = \tau(f_A)$ implies $a = \pm 1$, i.e. that f_A ia a translation. $\square$

Observe that if g is a parabolic isometry for the unit disc Δ with fixed point $\xi \in \partial\Delta$ then g is conjugated with the translation
$$h : z \longmapsto z + 1$$
acting on $\mathbb{H}$. This translation leaves the orthogonal lines $\{y = c\}$ fixed and shifts the geodesics ending at ∞ (i.e. the vertical lines) of a unit on the right; therefore the geodesics of Δ ending at ξ are permuted by g while the “horocycles” (that is circles tangent to $\partial\Delta$ at ξ) are invariant curves.

Theorem B.3.5. *An automorphism $g \neq id \in Aut(S^2)$ is elliptic if and only if $\tau(g) \in [0,4)$.*

Proof. If g is elliptic then is conjugated to a rotation f_ϑ with associated matrix

$$\begin{pmatrix} e^{i\vartheta} & 0 \\ 0 & e^{-i\vartheta} \end{pmatrix}$$

and hence $\tau(g) = \tau(f_\vartheta) = 4\cos^2\vartheta \in [0,4)$. Conversely if $\tau(g) \in [0,4)$ then there exists $\vartheta \in (0,\pi)$ such that $\tau(g) = 4\cos^2\vartheta$; moreover g is not parabolic and this implies that g has two fixed points. Thus g is conjugated to an isometry f with 0 and ∞ as fixed points; hence f has

$$\begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix}$$

as associated matrix for a suitable $a \in \mathbb{C}\backslash\{0\}$ and thus the equality $4\cos^2\vartheta = \tau(g) = \tau(f) = \big(a+\frac{1}{a}\big)^2$ implies that either $a = e^{i\vartheta}$ or $a = e^{-i\vartheta}$, that is f is a (non trivial) rotation. $\square$

Theorem B.3.6. *An automorphism $g \neq id$ is hyperbolic if and only if $\tau(g) > 4$.*

Proof. If g is hyperbolic then it is conjugated to an homothety f_k with associated matrix

$$\begin{pmatrix} \sqrt{k} & 0 \\ 0 & \frac{1}{\sqrt{k}} \end{pmatrix}$$

and hence $\tau(g) = \tau(f_k) = \big(\sqrt{k} + \frac{1}{\sqrt{k}}\big)^2 > 4$ because $k \neq 1$. Conversely if $\tau(g) > 4$ then

$$\tau(g) \;=\; \left(\sqrt{k} + \frac{1}{\sqrt{k}}\right)^2$$

for a suitable $k \neq 1$. Since $\tau(g) > 4$, g is not parabolic (neither elliptic) and it has exactly two fixed points; moreover g is conjugated to $f_k : z \longmapsto kz$ because $\tau(g) = \tau(f_k)$. $\square$

Notice that elliptic and hyperbolic automorphisms have two fixed points because rotations and homotheties do; hence the number of fixed points does not determine elliptic (and hyperbolic) automorphisms, in contrast of what happens for the parabolic ones. If g is elliptic for the unit disc Δ with fixed point ξ_1, ξ_2 then g is conjugate to a rotation h on S^2 (that leaves 0 and ∞ fixed); since the lines through the origin are twisted by h while the circles with center at the origin are preserved (the lines through the origin are the circles through 0 and ∞ while the circles with center at the origin have 0 and ∞ as inverse points) we have that g permutes the circles through ξ_1 and ξ_2 and leaves invariant the circles having ξ_1 and ξ_2 as inverse points. Instead if g is hyperbolic for Δ with ξ_1, ξ_2 fixed points then it is conjugated to an homothety h; since h leaves the lines through the origin fixed and permutes the circles with center in the origin we have that the circles through ξ_1 and ξ_2 in Δ are invariant under g while the circles that have ξ_1 and ξ_2 as inverse points are permuted.

B.4 Vector bundles

Definition B.4.1. *Let M be a n-dimensional manifold; a* ***vector bundle*** *of rank r on M is a smooth serjective function $\pi : E \longrightarrow M$, where E is a manifold (the total space), such that:*

1. *For all $p \in M$ the fiber $E_p := \pi^{-1}(p)$ is a r-dimensional vector space over $\mathbb{R}$;*
2. *For all $p \in M$ there exists a neighborhood U of p and a diffeomorphism $\chi : \pi^{-1}(U) \longrightarrow U \times \mathbb{R}^r$, called local trivialization, such that the diagram*

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\chi"] \arrow[dr, "\pi"'] & & U \times \mathbb{R}^r \arrow[dl, "p_1"] \\ & U & \end{tikzcd}$$

is commutative and for all $q \in U$ the map $\chi|_{E_q} : E_q \to \{q\} \times \mathbb{R}^r$ is a vector spaces isomrphism.

Remark. *The tangent bundle TM together with the projection $\pi : TM \longrightarrow M$ defined by $\pi(T_pM) := p$ is a vector bundle; analogously the cotangent bundle T^*M is a vector bundle.*

Definition B.4.2. *Given a vector bundle $\pi : E \longrightarrow M$, a **section** is a smooth function $\sigma : M \longrightarrow E$ such that $\pi \circ \sigma = id_M$ (i.e. $\sigma(p) \in E_p$ for all $p \in M$). The map $p \longmapsto O_p$, where $O_p \in E_p$ is the zero vector, is called zero section while a **local section** on $U \subseteq M$ is a section of $E_U = \pi^{-1}(U)$.*

In other words a section is a smooth way to associate at each $p \in M$ a vector in E_p; observe that the projection $\pi : M \times \mathbb{R}^r \longrightarrow M$ is obviously a vector bundle (called *trivial bundle*). In this case a section σ is of the form $\sigma(p) = (p, f(p))$ where $f : M \longrightarrow \mathbb{R}^r$ is a smooth function, i.e. the space of sections of the trivial bundle is canonically identified with $C^\infty(M, \mathbb{R}^r)$. One can prove that for a smooth manifold there always exists a non trivial section; this statement is no longer true if we suppose that the manifold is analitic or complex. The reader should refer to [**?**] or [8] for further details. In differential geometry are very important at least three particular spaces of sections:

- The space of section of the tangent bundle, denoted by $\mathcal{T}(M)$, whose elements are called *vector fields*; hence a vector field is a C^∞-way to associate at each $p \in M$ a vector in the tangent space T_pM.
- The space of section of the cotangent bundle, denoted by $A^1(M)$, whose elements are called *1-forms*; thus a 1-form is a smooth way to associate at each $p \in M$ a linear map $\omega : T_pM \longrightarrow \mathbb{R}$, i.e. an element in T_p^*M.
- The space sections of the k-forms bundle, denoted by $A^k(M)$, whose elements are called *k-forms*; as above a k-form is a C^∞-way to associate at each $p \in M$ a k-linear antisymmetric map $w : T_pM \times ... \times T_pM \longrightarrow \mathbb{R}$.

B.5 The Lie derivative

Definition B.5.1. *Given a vector field X an **integral curve** of X is a curve $\sigma : I \longrightarrow M$ such that $\sigma'(t) = X(\sigma(t))$ for all $t \in I$; moreover if $0 \in I$ and $\sigma(0) = p$ we say that σ is outgoing from p.*

The following theorem is a classical and massive result in differential geometry and goes under the name of *Existence and uniqueness of the local flow.* We do not give a proof of this theorem; one can refer for example to [2] for a complete proof.

Theorem B.5.2 (Existence and uniqueness of the local flow)**.** *Given a vector field X there exist only one open neighborhood U of $\{0\} \times M$ in $\mathbb{R} \times M$ and only one smooth function $\varphi : U \longrightarrow M$ (called **local flow** of X) such that:*

1. *For all $p \in M$ the set $U^p := \{t \in \mathbb{R} \mid (t, p) \in U\}$ is an open interval in $\mathbb{R}$ containing zero;*
2. *For all $p \in M$ the map $\varphi^p : U^p \longrightarrow M$ defined by $\varphi^p(t) := \varphi(t, p)$ is the maximal integral curves of X outgoing from p;*
3. *For all $t \in \mathbb{R}$ the set $U_t := \{p \in M \mid (t, p) \in U\}$ is open in M;*
4. *The map $\varphi_t : U_t \longrightarrow M$ defined by $\varphi_t(p) := \varphi(t, p)$ satisfies the following properties:*
   (a) *If $p \in U_t$ then $p \in U_{s+t}$ if and only if $\varphi_t(p) \in U_s$.*
   (b) *$\varphi_t(\varphi_s(p)) = \varphi_{t+s}(p)$, i.e. $t \longmapsto \varphi_t$ is a 1-parameter local group; in particular $\varphi_0 = id$ and $\varphi_t : U_t \longrightarrow U_{-t}$ is a diffemorphism with inverse φ_{-t}.*

Definition B.5.3. *A vector field X is called **complete** if its flow is defined on the whole space $\mathbb{R} \times M$, i.e. if each maximal integral curve has $\mathbb{R}$ as its domain of definition.*

Remark. *If M is compact then every vector field X is complete; if M is non compact a counterexample is given by*

$$M = \mathbb{R}^2 \setminus \{0\}, \qquad X = -\frac{p}{\|p\|}\,.$$

In this case $\varphi_t(p) = p - t\frac{p}{\|p\|}$ and hence the maximal integral curves are not defined on all $\mathbb{R}$.

Corollary B.5.4. *The following hold:*

1. $d(\varphi_t)_p(X) = X_{\varphi_t(p)}$;

2. *If* $f \in C^\infty(M)$ *and* $p \in M$ *then* $\frac{d}{dt}(f \circ \varphi^p)\Big|_{t=0} = (Xf)(p)$;

Denote by $d(\varphi_t)_p : T_pM \longrightarrow T_{\varphi_t(p)}M$ the differential of φ_t at p; thus $d(\varphi_t)_p$ is an isomorphism and

$$[d(\varphi_t)_p]^{-1} = d(\varphi_{-t})_{\varphi_t(p)} : T_{\varphi_t(p)}M \longrightarrow T_pM\,.$$

We say that a vector field Y is X-*invariant* if $d(\varphi_t)_pY = Y_{\varphi_t(p)}$ for all (t,p) in the domain of definition of the local flow φ; in other words Y is X-invariant if it is constant along the integral curves of X. Observe that if Y is X-invariant then $Y_p = d(\varphi_{-t})_{\varphi_t(p)}Y$; this fact leads to the following definition

Definition B.5.5. *Given two vector fields X and Y, the **Lie derivative** of Y along X is*

$$\mathcal{L}_XY(p) := \lim_{t\to 0} \frac{d(\varphi_{-t})_{\varphi_t(p)}Y - Y_p}{t}$$

where φ is the local flow of X.

One can extend the definition of *Lie derivative* along the vector field X to a generic k-form ω as follows; given a k-form ω and a vector field X the Lie derivative of ω along X is the k-form $\mathcal{L}_X\omega$ defined by

$$(\mathcal{L}_X\omega)(p) := \frac{d}{dt}\Big[\varphi_t^*\omega(p)\Big]_{t=0} \qquad \forall\, p \in M$$

where as usual φ_t is the local flow of X.

Proposition B.5.6. *Let M be a smooth manifold, then the following hold:*

1. *If $\{X_t\}$, $\{\omega_t\}$ are 1-parameter families of vector fields, resp. k-forms and φ_t is the local flow of X_t, i.e. the solution of the non-autonomous system*

$$\begin{cases} \frac{d}{dt}\varphi_t = X_t(\varphi_t); \\ \\ \varphi_0 = id; \end{cases}$$

   then $\frac{d}{dt}\Big(\varphi_t^*\omega_t\Big) = \varphi_t^*\Big(\mathcal{L}_{X_t}\omega_t + \frac{d}{dt}\omega_t\Big)$.

2. *If f is a smooth function defined on M then* $\mathcal{L}_Xf = df[X]$.

3. $\mathcal{L}_X(\alpha\wedge\beta) = \mathcal{L}_X\alpha\wedge\beta + \alpha\wedge\mathcal{L}_X\beta$.

4. $\mathcal{L}_X\omega = d\imath_X\omega + \imath_Xd\omega$ *(Cartan's idenity).*

B.6 The exponential map

Definition B.6.1. *Let M be a smooth manifold; a **Riemannian metric** on M is a smooth function $p \longmapsto g_p$ where g_p is a positive definite scalar product on T_pM for each $p \in M$. A **Riemannian manifold** is a pair (M,g) where M is a smooth manifold and g is a Riemannian metric on M.*

Hereafter we always suppose that the manifold M is endowed with a Riemannian metric g; thus the norm of a tangent vector and the angle between two tangent vectors in T_pM are well defined. If $\gamma : I \longrightarrow M$ is a differentiable curve then the *length* of γ is

$$L(\gamma) := \int_I |\dot\gamma(t)|\,dt\,.$$

If $p,q \in M$ are two points then the distance $d(p,q)$ between p and q is defined as the infimum of the length of curves joining p and q, that is

$$d(p,q) := \inf \{L(\gamma) \mid \gamma : [0,1] \longrightarrow M\,,\ \gamma(0) = p\,,\ \gamma(1) = q\}\,.$$

The manifold M endowed with the distance $d(\cdot,\cdot)$ is a metric space; moreover d induces on M a topology that is equivalent to the topology of manifold. In this context a *geodesic* is a curve $\gamma : I \longrightarrow M$ that locally realizes the distance; a geodesic might be not globally minimizing as one can see computing the geodesics of the sphere S^n. We say that a geodesic $\gamma : I \longrightarrow M$ is *maximal* if it can not be extended to a geodesic defined on an interval $J \supset I$; we have the following

Theorem B.6.2. *Let $p \in M$ and $v \in T_pM$, then there is only one maximal geodesic $\gamma_v : I_v \longrightarrow M$ such that $\gamma_v(0) = p$ and $\dot\gamma_v(0) = v$; moreover I_v is an open interval containing* 0.

Let $p \in M$ be a point and let $U_p \subseteq T_pM$ the subset of the tangent space of the vectors $v \in T_pM$ such that I_v (as above) contains the point 1; then the *exponential map* in p is

$$\exp_p : U_p \longrightarrow M\,, \qquad \exp_p(v) := \gamma_v(1)\,.$$

Proposition B.6.3. *U_p is open and contains the origin; moreover the differential of the exponential map in the origin is the identity and hence $\exp_p$ is a local diffeomorphism near the origin.*

Given a point $p \in M$ the *injectivity radius* of M in p is defined as

$$\mathrm{inj}_p\,(M) := \sup \{r > 0 \mid \exp_p|_{B_0(r)} \text{ is a diffeomorphism}\}$$

where $B_0(r)$ denotes the ball of radius r centered at the origin in T_pM. If $r < \mathrm{inj}_p\,(M)$ then $\exp_p$ maps the ball $B_0(r) \subseteq T_pM$ in the ball of radius r and center p in M, i.e.

$$\exp_p\,(B_0(r)) \;=\; B_p(r)\,;$$

therefore $B_p(r)$ is effectively diffeomorphic to an open ball in $\mathbb{R}^n$. If R is big this might be no longer true, indeed for instance if M is compact then for r large enough we have $B_p(r) = M$; the *injectivity radius* of M is the infimum of the injectivity radius of M in p, that is

$$\mathrm{inj}\,(M) := \inf_{p \in M}\ \mathrm{inj}_p\,(M)\,.$$

Proposition B.6.4. *If M is compact then $\mathrm{inj}\,(M) > 0$.*

Proposition B.6.5. *A curve with length $< 2 \cdot \mathrm{inj}\,(M)$ is omotopically equivalent to a point.*

Proof. Let $x \in \gamma(I)$; since $L(\gamma) < 2 \cdot \mathrm{inj}\,(M)$ the curve is completely in $B_x(r)$ for a suitable $r < \mathrm{inj}\,(M) \le \mathrm{inj}_x\,(M)$; therefore $B_x(r)$ is diffeomorphic to a ball in $\mathbb{R}^n$ which is contractile. □

Theorem B.6.6 (Hopf-Rinow)**.** *Let M be connected, then the following are equivalent:*

1. *M is complete.*
2. *A subset $U \subseteq M$ is compact if and only if it is closed and bounded.*
3. *Each maximal geodesic is defined on all $\mathbb{R}$.*
4. *For each $p \in M$ the exponential map $\exp_p$ is defined on the whole tangent space T_pM.*

Corollary B.6.7. *If M is compact then M is complete.*

B.7 Sectional curvature

We shall not introduce the concept of curvature in the most general way because it turn to be a tool too sophisticated for our goal; indeed in hyperbolic geometry all the manifolds will have constant curvature and the Riemann tensor is not necessary. In dimension 2, that is if M is a surface, then all the notions of curvature are reduced to the concept of *Gaussian curvature*; if $M \subseteq \mathbb{R}^3$ the Gaussian curvature is the product of the principal curvatures, but in general the principal curvatures are not defined. Thus we need to proceed in another way; if M is a surface and $p \in M$ then there is $\epsilon > 0$ such that the ball $B_p(\epsilon)$ is diffeomorphic to a ball in $\mathbb{R}^2$, although its volume is in general different from the volume of an Euclidean ball. Hence we define the Gaussian curvature of M in p as

$$K := \lim_{\epsilon \to 0} \left[\left(\pi \epsilon^2 - \mathrm{Vol}\left(B_p(\epsilon) \right) \right) \frac{12}{\pi \epsilon^4} \right] ;$$

notice that K is positive (resp. negative) if $B_p(\epsilon)$ has area lower (resp. bigger) than the Euclidean one. In general if M is a manifold and $p \in M$ is a point we consider a plane $W \subseteq T_pM$ and define the *sectional curvature* of M along (p, W) to be the Gaussian curvature of the surface $S := \exp_p(U_p \cap W)$, where $U_p \subseteq T_pM$ is such that the exponential map restricted to U_p is a diffeomorphism.

Definition B.7.1. *We say that M has* ***constant sectional curvature*** *K if the sectional curvature associated to each $p \in M$ and to each $W \subseteq T_pM$ is equal to K.*

Observe that if M has constant sectional curvature then, up to rescaling the metric g on M we can suppose that M has constant sectional curvature equal to -1, 0 or 1; the models of this manifold will be respectively the hyperbolic space $\mathbb{H}^n$, the space $\mathbb{R}^n$ and the sphere S^n.

B.8 Group actions on topological spaces

Let G be a group and let X be a topological space (usually G is a group of homeomorphisms of X), then we say that G acts *freely* on X if $g(x) \neq x$ for all $x \in X$ and for all $g \neq id$. Furthermore we say that the action of G is *properly discontinuous* if for all $x, y \in X$ there exist U_x, U_y open neighbourhoods of x, y such that

$$g(U_x) \cap U_y \neq \emptyset$$

only for a finite number of $g \in G$. Hereafter all the spaces are supposed to be "decent", that is Hausdorff, connected and locally connected by arcs.

Proposition B.8.1. *Given a topological space X and a group G, the following are equivalent:*

1. *The action of G is free and properly discontinuous.*
2. *X/G is Hausdorff and the projection $p : X \longrightarrow X/G$ is a covering.*

Let $p : E \longrightarrow X$ be a covering; an *automorphism* of covering is a homeomorphism $f : E \longrightarrow E$ such that $p \circ f = p$, that is such that the following diagram

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "p"'] & & E \arrow[dl, "p"] \\ & X & \end{tikzcd}$$

is commutative. The set of automorphisms of $p : E \longrightarrow X$ has a natural structure of group and it is denoted by $\mathrm{Aut}\,(p)$; moreover the action of $\mathrm{Aut}\,(p)$ on E is free and properly discontinuous. Thus the following question arises naturally: the covering $p : E \longrightarrow X$ is the quotient of E with respect to this action or not? The answer to this question in general is negative, but however it turns to be positive for a large class of coverings, the so called *regular coverings.* We say that a covering $p : E \longrightarrow X$ is regular if the canonical homomorphism of group induced by p maps the fundamental group of E into a normal subgroup of $\pi_1(X)$, i.e. if

$$p_*(\pi_1(E)) \lhd \pi_1(X)\,;$$

in this case it is possible to prove that $X = E/\mathrm{Aut}\,(p)$ and

$$\mathrm{Aut}\,(p) \cong \pi_1(X)\Big/p_*(\pi_1(E))\,.$$

Corollary B.8.2. *If $p : \widetilde{X} \longrightarrow X$ is a universal covering then $X = \widetilde{X}/Aut\,(p)$ and*

$$Aut\,(p) \cong \pi_1(X)\,.$$

If M is a Riemannian manifold then the *Myers-Steenrod theorem* states that the set $\mathrm{Isom}\,(M)$ of isometries of M has a natural structure of Lie group with respect to which the map

$$\mathrm{Isom}\,(M) \times M \longrightarrow M \times M\,, \qquad (\varphi, p) \longmapsto (\varphi(p), p)$$

is smooth and proper (i.e. the inverse image of a compact is compact).

Proposition B.8.3. *Let $G < Isom\,(M)$ be a subgroup of isometries of a Riemannian manifold, then the action of G is properly discontinuous if and only if G is discrete in $Isom\,(M)$.*

B.9 Hyperbolic manifolds

The hyperbolic space $\mathbb{H}^n$ can be equivalently defined as:

- The *hyperboloid* $I^n = \{(x', x_{n+1}) \in \mathbb{R}^{n+1} \mid \|x'\|^2 - |x_{n+1}|^2 = -1\}$ endowed with the metric

$$g := dx_1 \otimes dx_1 + \ \dots\ + dx_n \otimes dx_n - dx_{n+1} \otimes dx_{n+1}\,.$$

- The *Poincaré ball* $D^n = \{x \in \mathbb{R}^n \mid \|x\| < 1\}$ considered with the metric

$$g := \left(\frac{2}{1 - \|x\|}\right)^2 \cdot (dx_1 \otimes dx_1 + \ \dots\ + dx_n \otimes dx_n)\,.$$

- The *upper half-sace* $H^n = \{x \in \mathbb{R}^n \mid x_n > 0\}$ with the metric

$$g := \frac{1}{x_n^2} \cdot (dx_1 \otimes dx_1 + \ \dots\ + dx_n \otimes dx_n)\,.$$

These three models are isometrically equivalent and will be denoted indistinctly by $\mathbb{H}^n$; a simple computation shows that the geodesics of $\mathbb{H}^n$ are defined on all $\mathbb{R}$ and hence by the Hopf-Rinow theorem the hyperbolic space is complete. The isometries of $\mathbb{H}^n$ can be classified by their number of fixed points, in the sense that for any isometry φ of the hyperbolic space one and only one of the following facts hold:

1. φ has at least one fixed point in $\mathbb{H}^n$ (in this case we say that φ is *elliptic*).
2. φ has no fixed points in $\mathbb{H}^n$ and exactly one in $\partial\mathbb{H}^n$ (in this case we say that φ is *parabolic*).
3. φ has no fixed points in $\mathbb{H}^n$ and exactly two in $\partial\mathbb{H}^n$ (in this case we say that φ is *hyperbolic*).

One can prove that in $\mathbb{H}^2$ the disc of radius r centered at the origin has area

$$A(r) = 2\pi(\cosh r - 1)\,;$$

hence it follows immediately that the hyperbolic space $\mathbb{H}^n$ has constant sectional curvature -1.

Definition B.9.1. *A **hyperbolic manifold** is a Riemannian manifold in which every point p has an open neighbourhood isometric to an open subset of $\mathbb{H}^n$.*

If M is an hyperbolic manifold then it has constant sectional curvature -1; it is possible to prove that also the converse is true, i.e. that if a manifold M has constant sectional curvature -1 then M is hyperbolic. Recall that if $p : \widetilde{M} \longrightarrow M$ is a covering between topological spaces and M is a manifold (resp. Riemannian manifold) then there is only one structure of manifold (resp. Riemannian manifold) on $\widetilde{M}$ such that p is a local diffeomorphism (resp. isometry); in other words we can always "push up" the structure of manifold (resp. Riemannian manifold) along coverings and this can be made in a unique way. Conversely in general is not possible to "push down" the structure of manifold (or Riemannian manifold) along coverings; however this can (sometimes) be made if G is a group acting freely and discontinuously on a manifold M.

Proposition B.9.2. *Let G be a group acting freely and discontinuously on a manifold (resp. Riemannian manifold) M and let $p : M \longrightarrow M/G$ be the induced covering; then there is a structure of manifold (resp. Riemannian manifold) on M/G such that the map p is a local diffeomorphism (resp. isometry) if and only if G is a group of diffeomorphisms (resp. isometries) of M. Moreover, if it exists, this structure is unique.*

It follows immediately from the proposition above and from proposition B.8.3 that if $G < \mathrm{Isom}\,(M)$ is a discrete subgroup of isometries of M acting freely on M then M/G is a manifold and the map $p : M \longrightarrow M/G$ is a covering and a local isometry. Hence we get the following

Corollary B.9.3. *If $G < Isom\,(\mathbb{H}^n)$ is a descrete subgroup composed only by hyperbolic and parabolic isometries then the quotient M/G is a hyperbolic manifold.*

It is just a simple consequence of the Hopf-Rinow theorem the fact that if M, N are Riemannian manifold, with N complete, and $p : M \longrightarrow N$ is a covering then M is complete. Therefore if M is a non complete hyperbolic manifold its universal covering $\widetilde{M}$ will be a simply connected non complete hyperbolic manifold (for instance an open and simply connected subset of $\mathbb{H}^n$); however it is possible to show that it always exists a natural map

$$D : \widetilde{M} \longrightarrow \mathbb{H}^n\,.$$

Moreover if $\widetilde{M}$ is complete then the map D is an isometry; using all these facts one can finally prove the following important theorem about hyperbolic complete manifolds.

Theorem B.9.4. *A complete Riemannian manifold is hyperbolic if and only if it is isometric to $\mathbb{H}^n/G$ for a suitable discrete group of parabolic and hyperbolic isometries G.*

Proof. If G is descrete and contains only parabolic and hyperbolic isometries then $\mathbb{H}^n/G$ is a complete hyperbolic manifold. Conversely, let M be a complete hyperbolic manifold; then its universal covering $\widetilde{M}$ is a complete hyperbolic manifold and hence it is isometric to $\mathbb{H}^n$. Therefore $M = \mathbb{H}^n/\mathrm{Aut}\,(p)$ and $\mathrm{Aut}\,(p)$ must be a descrete group of isometries without fixed points. □

Let $G < \mathrm{Isom}\,(\mathbb{H}^n)$ be discrete without elliptic elements; it is possible to prove that if G contains a parabolic element then the quotient $\mathbb{H}^n/G$ has injectivity radius 0 and hence $\mathbb{H}^n/G$ can not be compact because compact manifolds have positive injectivity radius. Thus we get the following

Corollary B.9.5. *If $\mathbb{H}^n/G$ is compact then G contains only hyperbolic elements.*

Bibliography

[1] J. Mather, *Action minimizing invariant measures for positive definite lagrangian systems*, Math. Z 207, 1991.

[2] M. do Carmo, *Differential geometry of curves and surfaces*, 1976.

[3] M.Abate, F.Tovena, *Geometria differenziale*, Springer, 2011.

[4] R. Tyrrell Rockafellar, *Convex Analysis*, Princeton Mathematical series, 21, Princeton University Press, New Jersey, 1970.

[5] J.M. Ball and V.J. Mizel, *One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation*, Arch. Rational Mech. Anal., 1985.

[6] G. Contreras and R. Iturriaga, *Global minimizers of autonomous lagrangians*, Cimat, Mexico, 2000.

[7] A. Sorrentino, *Lecture notes on Mather's theory for lagrangian systems.*

[8] E. Kreyszig, *Differential geometry*, 1991.

[9] R. Mañé, *Lagrangian flows: the dynamics of globally minimizing orbits*, International conference on dynamical systems (Montevideo,1995).

[10] G. Paternain and M. Paternain, *Critical values of autonomous lagrangian systems*, Comment. Math. Helv. (1997).

[11] N. Kryloff and N. Bogoliouboff, *La théorie générale de la measure dans son application á l'étude des systémes dynamiques de la mécanique non linéaire*, Ann. of Math. , 1937.

[12] D. McDuff, D. Salamon, *Introduction to symplectic topology*, Oxford Mathematical monographs, 1998.

[13] A. Fathi, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sér. I Math., 1998.

[14] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, New York, 1992.

[15] R. Mañé, *On the minimizing measures of Lagrangian dynamical systems*, Nonlinearity, 1992.

[16] H. Cartan, *Theorie elementaire des fonctions analytiques d'une ou plusieurs variables complexes*, Coll. ens. de sciences, 1961.

[17] R. Bowen, *Symbolic dynamics for hyperbolic flows*, Amer. J. Math., 1973.

[18] R. Adler, B. Weiss, *Entropy, a complete metric invariant for automorphisms of the Torus*, Proc. Nat. Acad. Sci. U.S.A, 1967.

[19] Ya. G. Sinai, *Markovian partitions and Y-diffeomorphisms*, Func. Anal. Appl., 1968.

[20] T. Bedford, M. Keane, C. Series, *Ergodic theory, symbolic dynamics and hyperbolic spaces*, Oxford Sci. Pub., 1991.

[21] G. Hedlund, *Fuchsian groups and transitive horocycles*, Duke Math. J., 1936.

[22] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer, 1983.

[23] K. Burns, G. Paternain, *Anosov magnetic flows, critical value and topological entropy*, 2002.

[24] B. Martelli, *Lezioni di geometria iperbolica.*

[25] R. Benedetti, C. Petronio, *Lectures on hyperbolic geometry*, Springer, 1992.

[26] A. Abbondandolo, *Lecture notes on Floer homology.*

[27] Bernard, *The Lax-Oleinik semigroup, a Hamiltonian point of view.*

[28] G. Contreras, R. Iturriaga, G. Paternain, *Lagrangian graphs, mininizing measures and Mañé's critical values*, Geom. Funct. Anal., 1998.

[29] R. Iturriaga, H.S. Morgado, *Finsler metric and action potential*, Proc. Amer. Math. Soc, 2000.

[30] V. I. Arnold, *Some remarks on flows of line elements and frames*, Sov. Math. Dokl., 1961.

[31] D.V. Anosov, Y.G. Sinai, *Certain smooth ergodic systems*, Uspehi Mat. Nauk, 1967.