# HERGLOTZ' VARIATIONAL PRINCIPLE AND LAX-OLEINIK EVOLUTION

PIERMARCO CANNARSA, WEI CHENG, LIANG JIN, KAIZHI WANG, AND JUN YAN

ABSTRACT. We develop an elementary method to give a Lipschitz estimate for the minimizers in the problem of Herglotz' variational principle proposed in [17] in the time-dependent case. We deduce Erdmann's condition and the Euler-Lagrange equation separately under different sets of assumptions, by using a generalized du Bois-Reymond lemma. As an application, we obtain a representation formula for the viscosity solution of the Cauchy problem for the Hamilton-Jacobi equation

$$D_t u(t,x) + H(t,x,D_x u(t,x),u(t,x)) = 0$$

and study the related Lax-Oleinik evolution.

## 1. INTRODUCTION

1.1. **Introduction.** A basic problem of calculus of variations is to minimize the action functional

$$\int_a^b L(s,\xi(s),\dot{\xi}(s))\,ds$$

over the set of absolutely continuous curves $\xi$ connecting two points $x,y\in\mathbb{R}^n$. It has been studied up to now for almost three hundred years. Beyond the issue of the existence of minimizers, much of the attention in the calculus of variations has been devoted to necessary conditions for optimality. Another essential point of the analysis is the Lipschitz regularity of minimizers. This property has many applications, for instance to Euler-Lagrange equations, where it can be used to exclude the Lavrentiev phenomenon (see, for instance, [10] for a survey on this topic). The Lipschitz regularity of minimizers is the subject of an extensive literature (see, for instance, [23, 1, 52, 25, 21, 11, 22]).

This paper is devoted to the *generalized variational principle* proposed by Gustav Herglotz in 1930 ([36, 37]). Such a result generalizes classical variational principles by defining a functional whose extrema are sought by a differential equation.

More precisely, let $L\in C^2(\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R},\mathbb{R})$ and $\xi:[a,b]\to\mathbb{R}^n$ be any piecewise $C^1$ curve. The functional $u_\xi$ is defined in an implicit way by the ordinary differential equation

$$\dot{u}_\xi(s) = L(s,\xi(s),\dot{\xi}(s),u_\xi(s)),\quad s\in[a,b], \tag{1.1}$$

with $u_\xi(a)=u\in\mathbb{R}$, for $b>a$. The so-called Herglotz' variational principle is to seek an extremal $\xi$ of the functional

$$u[\xi] := u_\xi(b) - u = \int_a^b L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds,$$

*Date*: February 25, 2020.
2010 *Mathematics Subject Classification.* 35F21, 49L25, 37J50.
*Key words and phrases.* Herglotz' variational principle, Hamilton-Jacobi equation, viscosity solution.

where $u_\xi$ is determined by (1.1). We call $\xi$ is an extremal of $u[\xi]$ if $\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}u[\xi+\varepsilon\eta]=0$ for arbitray piecewise $C^1$ curve $\eta$ such that $\eta(a)=\eta(b)=0$. Herglotz' variational principe gurantees that any $C^2$ extremal of the functional $u[\xi]$ must satisfy the so-called Herglotz equation

$$\frac{d}{ds}L_v = L_x + L_uL_v. \tag{1.2}$$

Herglotz reached the idea of the generalized variational principle through his work on contact transformations and their connections with Hamiltonian systems and Poisson brackets. The reader can find more information on the problem and its rather wide connections in [17] (see also [35, 32, 33]) and the references therein. However, to our knowledge, there is no rigorous approach to this problem in a modern setting including the existence and regularity results.

1.2. **Assumptions on $L$.** Now, we impose our assumptions on the Lagrangian $L$. Let $L=L(t,x,v,r):\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ be a function of class $C^1$ such that the following standing assumptions are satisfied:

(L1) $L(t,x,\cdot,r)$ is strictly convex for all $(t,x,r)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}$.

(L2) There exist two superlinear functions $\overline{\theta}_0,\theta_0:[0,+\infty)\to[0,+\infty)$ and two $L^\infty_{\text{loc}}$-functions $c_0,c_1:\mathbb{R}\to[0,+\infty)$, such that

$$\overline{\theta}_0(|v|)+c_1(t)\geqslant L(t,x,v,0)\geqslant\theta_0(|v|)-c_0(t),\quad (t,x,v)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n.$$

(L3) There exists an $L^\infty_{\text{loc}}$-function $K:\mathbb{R}\to[0,+\infty)$ such that

$$|L_r(t,x,v,r)|\leqslant K(t),\quad (t,x,v,r)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}.$$

(L4) There exists two $L^\infty_{\text{loc}}$-functions $C_1,C_2:\mathbb{R}\to[0,\infty)$ such that

$$|L_t(t,x,v,r)|\leqslant C_1(t)+C_2(t)L(t,x,v,r),\quad (t,x,v,r)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}.$$

There are various conditions that may replace (L4). We will mainly focus on the following substitution of (L4):

(L4') There exist two $L^\infty_{\text{loc}}$-functions $C_1,C_2:\mathbb{R}\to[0,\infty)$ such that for all $(t,x,v,r)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}$

$$\max\{|L_x(t,x,v,r)|,|L_v(t,x,v,r)|\}\leqslant C_1(t)+C_2(t)L(t,x,v,r).$$

*Remark* 1.1. If $a<b$ are fixed and $L$ is restricted on $[a,b]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}$, then the $L^\infty_{\text{loc}}$-functions $c_0(t),c_1(t),K(t),C_1(t),C_2(t)$ appear in our assumptions on $L$ can be chosen as constants, say $c_0,c_1,K,C_1,C_2$ (we also set $c_1=0$ for convenience). In fact, we can also assume $C_1\in L^1$ in condition (L4) and (L4') respectively.

1.3. **Herglotz' variational principle.** Fix $x,y\in\mathbb{R}^n$, $a<b$ and $u\in\mathbb{R}$. Set

$$\Gamma^{a,b}_{x,y}=\{\xi\in W^{1,1}([a,b],\mathbb{R}^n):\xi(a)=x,\ \xi(b)=y\}.$$

For any given $\xi\in\Gamma^{a,b}_{x,y}$, we consider the Carathéodory equation

$$\begin{cases}\dot u_\xi(s)=L(s,\xi(s),\dot\xi(s),u_\xi(s)), & a.e.\ s\in[a,b],\\ u_\xi(a)=u.\end{cases} \tag{1.3}$$

We define the action functional

$$J(\xi):=\int_a^b L(s,\xi(s),\dot\xi(s),u_\xi(s))\,ds, \tag{1.4}$$

where $\xi \in \Gamma^{a,b}_{x,y}$ and $u_\xi$ is uniquely determined by (1.3) and Proposition A.1. Our purpose is to minimize $J(\xi)$ over

$$\mathcal{A} = \mathcal{A}^{a,b,u}_{x,y} = \{\xi \in \Gamma^{a,b}_{x,y} : \text{(1.3) admits an absolutely continuous solution } u_\xi\}.$$

Notice that $\mathcal{A} \neq \varnothing$ because it contains all piecewise $C^1$ curves connecting $x$ to $y$. It is not hard to check that, for each $r \in \mathbb{R}$,

$$\mathcal{A} = \mathcal{A}' := \{\xi \in \Gamma^{a,b}_{x,y} : s \mapsto L(s, \xi(s), \dot{\xi}(s), r) \text{ belongs to } L^1([a,b])\}.$$

In fact, what we are studying is a variational problem under a very special non-holonomic constraint. The readers can refer to, for instance, [33]. Our work is essentially motivated by the recent works [53, 17, 56].

*Remark* 1.2. If we replace the same Lagrangian $L$ in both the action functional defined in (1.4) and the constraint in (1.3) by two distinct Lagrangians $L_1$ and $L_2$, the original Herglotz' problem become a much more difficult variational problem with non-holonomic constraint. This is a possible way to understand the variational nature of the relevant problem of Hamilton-Jacobi systems (see, for instance, [27, 47, 43, 39, 40]).

**Proposition 1.3.** *Fix $x, y \in \mathbb{R}^n$, $b > a$ and $u \in \mathbb{R}$. Under conditions* (L1)-(L3)*, the functional*

$$\mathcal{A} \ni \xi \mapsto J(\xi) = \int_a^b L(s, \xi(s), \dot{\xi}(s), u_\xi(s))\, ds,$$

*where $u_\xi$ is determined by* (1.3)*, admits a minimizer.*

The proof of Proposition 1.3 is given in Appendix B (see [17] for the time-independent case) under relaxed conditions (L1°)-(L3°).

1.4. **Erdmann condition and Herglotz equation.** From the technical point of view, this is the main part of this paper. Since the action functional $J$ is essentially defined in an implicit way, to our knowledge, all the methods in the standard references such as [23], [1] or [25] can not be applied directly. In the previous paper [17], due to summability issues, we solved this problem under restrictive growth conditions on $L$ for the autonomous case. In this paper, appealing to additional technical tools, we solve this problem as follows:

(1) We improve the classical du Bois-Reymond lemma in the calculus of variations proving that such a lemma holds even if the test functions are selected in a restricted space. More precisely, suppose $f, g \in L^1([a,b])$, $\delta \in L^\infty([a,b])$ and $\delta(s) > 0$ for almost all $s \in [a,b]$. Set the family of test functions as $\Omega = \{\beta \in L^\infty([a,b]) : \int_a^b \beta(s)\, ds = 0, |\beta| \leqslant \delta, a.e.\}$. We will show, if

$$\int_a^b f(s) b_\beta(s) + g(s)\beta(s)\, ds = 0, \quad \beta \in \Omega,$$

where $b_\beta(s) := \int_a^s \beta(r)\, dr$ for $\beta \in \Omega$, then there exists a continuous representative $\tilde{g}$ of $g$ such that $\tilde{g}$ is absolutely continuous on $[a,b]$ and $\tilde{g}'(s) = f(s)$ for almost all $s \in [a,b]$.

(2) We have to deal with the problem under various sets of conditions separately. If condition (L1)-(L3) together with (L4) are satisfied, we will adopt the method of [1] based on reparameterization. Without loss of generality we set $[a,b] = [0,t]$ for $t > 0$. For any measurable function $\alpha : [0,t] \to [1/2, 3/2]$ satisfying $\int_0^t \alpha(s)\, ds = t$, we define $\tau(s) = \int_0^s \alpha(r)\, dr$ for $s \in [0,t]$. Note that $\tau : [0,t] \to [0,t]$ is a bi-Lipschitz map.

Now, let $\xi \in \Gamma^{0,t}_{x,y}$ be a minimizer of $J$, and $\alpha \in \Omega$ as above. We define the reparameterization $\eta$ of $\xi$ by $\eta(\tau) = \xi(s(\tau))$ where $s(\tau)$ is the inverse of $\tau(s)$. It follows that $\dot\eta(\tau) = \dot\xi(s(\tau))/\alpha(s(\tau))$. Let $u_\eta$ be the unique solution of (1.3) with initial condition $u_\eta(0) = u$. Then we have that

$$\begin{aligned} J(\xi) \leqslant J(\eta) &= \int_0^t L(\tau, \eta(\tau), \dot\eta(\tau), u_\eta(\tau))\,d\tau \\ &= \int_0^t L(\tau(s), \xi(s), \dot\xi(s)/\alpha(s), u_{\xi,\alpha}(s))\alpha(s)\,ds \end{aligned}$$

where $u_{\xi,\alpha}$ solves

$$\dot u_{\xi,\alpha}(s) = L(\tau(s), \xi(s), u_{\xi,\alpha}(s), \dot\xi(s)/\alpha(s))\alpha(s), \quad u_{\xi,\alpha}(0) = u.$$

Define the functional $\Lambda : \Omega \to \mathbb{R}$ by

$$\Lambda(\alpha) = u_{\xi,\alpha}(t)$$

with $u_{\xi,\alpha}$ as above. We write $\alpha = 1 + \beta$. We should verify

$$0 = \frac{d}{d\varepsilon}\Lambda(1+\varepsilon\beta)|_{\varepsilon=0} = \int_a^b \left\{ E\cdot\beta - e^{-\int_a^s L_u\,dr} L_t(s,\xi,\dot\xi,u_\xi)\cdot b_\beta \right\}\,ds,$$

where $E(s) = e^{-\int_a^s L_u\,dr}\cdot\left\{ L_v(s,\xi(s),\dot\xi(s),u_\xi(s))\cdot\dot\xi(s) - L(s,\xi(s),\dot\xi(s),u_\xi(s))\right\}$, to obtain the Erdmann condition. From technical point of view, we need validate the convergence by using Lebesgue's theorem. That means we need check the required summability issues. This problem is solved by using our conditions (L1), (L3) and (L4) and the restriction of $\beta \in \Omega$. Now, invoking our generalized du Bois-Reymond lemma, we obtain the Erdmann condition

$$\frac{d}{ds}\,E(s) = -e^{-\int_a^s L_u\,dr} L_t(s), \quad a.e.\ s \in [0,t]. \tag{1.5}$$

(3) If conditions (L1)-(L3) together with (L4') are satisfied, we use the standard variation $\xi_\varepsilon = \xi + \varepsilon\eta$. Also for the summability difficulty, we restrict our $\eta \in \Omega$. One can deduce the Herglotz equation (1.2) on $[a,b]$ almost everywhere by using the generalized du Bois-Reymond lemma.

If $L$ is of class $C^2$, then any minimizer $\xi$ of (1.4) is as smooth as $L$ and $\xi$ satisfies Herglotz equation (1.2) on $[a,b]$ where $u_\xi$ is of class $C^2$ and satisfies Carathéodory ODE (1.3). Let $H$ be the associated Hamiltonian defined by

$$H(t,x,p,r) = \sup_{v\in\mathbb{R}^n}\{p\cdot v - L(t,x,v,r)\}, \quad t \in \mathbb{R}, (x,v) \in \mathbb{R}^n\times\mathbb{R}^n, r \in \mathbb{R}.$$

Then $H$ is also of class $C^2$ and satisfies certain standard conditions.

Set $p(s) = L_v(s,\xi(s),\dot\xi(s),u_\xi(s))$. Then the arc $(\xi, p, u_\xi)$ satisfies the following Lie equation

$$\begin{cases} \dot\xi = H_p(s,\xi,p,u_\xi), \\ \dot p = -H_x(s,\xi,p,u_\xi) - H_u(s,\xi,p,u_\xi)p, & s \in [a,b], \\ \dot u_\xi = p\cdot\dot\xi - H(s,\xi,p,u_\xi). \end{cases} \tag{1.6}$$

Equation (1.6) is a special kind of contact system. The readers can also recognize (1.6) as the system of characteristics. This system is widely studied in mathematics (see, for instance, [4, 28] for general information and [51, 55, 54, 46, 57, 20] especially on connections to Aubry-Mather theory and Hamilton-Jacobi equations), mechanics and mathematical physics (see, for instance, [7, 8, 45] and [49, 38, 50, 44] for Nosé-Hoover dynamics).

1.5. **Hamilton-Jacobi equations of contact type.** As an application, this paper establishes a connection between Herglotz' variational problem and the Hamilton-Jacobi equation

$$\text{(HJ)} \qquad \begin{cases} D_t u(t,x) + H(t,x,D_x u(t,x),u(t,x)) = 0 \\ \qquad\qquad\qquad\qquad\qquad\quad u(0,x) = \phi(x) \end{cases} \quad x \in \mathbb{R}^n, t > 0,$$

the solution of which is given by the related Lax-Oleinik evolution.

We suppose $L$ is of class $C^1$ satisfying conditions (L1)-(L3) together with (L4) (resp. (L4')), with $H$ being the associated Hamiltonian. Fix $x, y \in \mathbb{R}^n$, $t_2 > t_1$ and $u \in \mathbb{R}$. We define

$$h_L(t_1,t_2,x,y,u) := \inf_{\xi \in \mathcal{A}^{t_1,t_2}_{x,y}} \int_{t_1}^{t_2} L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\, ds,$$

where $u_\xi$ is determined by the associated Carathéodory equation. The function $h_L$ is called the fundamental solution of (HJ).

To study (HJ) for a wider class of the initial data. We suppose that $\phi$ is a real-valued function on $\mathbb{R}^n$ which is lower semi-continuous and $(\kappa_1, \kappa_2)$-Lipschitz in the large (see Definition 3.3). The main result is that

$$u(t,x) = \inf_{y \in \mathbb{R}^n} \{\phi(y) + h_L(0,t,y,x,\phi(y))\}$$

is finite-valued and it is a viscosity solution of (HJ). We also introduce the Lax-Oleinik evolution in this context and discuss the related dynamic programming principle. A systematic approach to this problem from Lagrangian formalism will be our task in the future.

*Example* 1.4. Let $V$ be a smooth real-valued function on $\mathbb{R}^n \times \mathbb{R}$, $\lambda \in \mathbb{R}$ and let

$$L(s,x,v,r) = L_0(s,x,v) - \lambda r,$$

where $L_0 = \frac{1}{2}|v|^2 - V(x,t)$. Then the associated Herglotz equation, i.e.,

$$\ddot{x} + \lambda \dot{x} + \nabla_x V(x,t) = 0,$$

is a Duffing-type equation, which is rather widely studied in many fields such as mechanics, nonlinear physics and engineering (see, for instance, [48]). Recall that the associated Hamiltonian has the form $H = H_0(s,x,p) + \lambda r$ where $H_0$ is the Fenchel-Legendre dual of $L_0$. This model is also closely related to discounted Hamilton-Jacobi equations in PDE and calculus of variations and optimal control [2, 3, 18, 34, 26, 41, 42, 57, 20].

The paper is organized as follows: In Section 2, we give a detailed proof of the Erdmann condition and Herglotz equation based on our generalized du Bois-Reymond lemma under various kind of conditions. Then we obtain the expected Lipschitz estimates. In Section 3, we apply Herglotz' variational principle to Hamilton-Jacobi equation (HJ). We have three appendices. In Appendix A, We collect useful material from analysis and differential equations. The Main part of Appendix B is composed of the details of the proofs of a Tonelli-like existence result and some necessary a priori estimates under a series of suitably weaker conditions. In Appendix C, we explain how to move Herglotz' variational principle to manifolds.

**Acknowledgement** All of the authors acknowledge the MIUR support from Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006. Personally, Piermarco Cannarsa was supported in part by the National Group for Mathematical Analysis, Probability and Applications (GNAMPA) of the Italian Istituto Nazionale di Alta Matematica Francesco Severi and by Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006. Wei Cheng was supported in part by National Natural Scientific Foundation of China (Grant No.11871267, No.11790272 and No.11631006). Liang Jin was supported in part by Natural Scientific Foundation of China (Grant No. 11901293 and No. 11571166) and Start-up Foundation of Nanjing University of Science and Technology (No. AE89991/114). Kaizhi Wang was supported in part by Natural Scientific Foundation of China (Grant No. 11771283 and No. 11931016). Jun Yan was supported in part by Natural Scientific Foundation of China (Grant No.11790272 and No.11631006). The authors are grateful to Qinbo Chen, Cui Chen, Jiahui Hong, Shengqing Hu and Kai Zhao for helpful discussions. The authors would like to thank the anonymous referees for their careful reading and useful comments on the original version of this paper, which have helped them to improve the presentation significantly.

## 2. NECESSARY CONDITIONS AND LIPSCHITZ ESTIMATES

The main purpose of this section is to give a Lipschitz estimate of any minimizer $\xi$ of (1.4) and to derive some necessary conditions such as the generalized Euler-Lagrange equation (Herglotz equation) and Erdmann condition. Unlike the autonomous case studied in [17], we will deal with the problem under various kind of conditions. It is worthing noting that

– one can deduce the Erdmann condition for the "energy function" directly under the conditions (L1)-(L3) together with (L4);
– one can also deduce the Herglotz equation directly under the conditions (L1)-(L3) together with (L4')

A key tool is the following lemma of du Bois-Reymond type (see Theorem 2.1). By using such a result, one can get the required Lipschitz estimate after having derived either Erdmann condition or Herglotz equation.

### 2.1. **A generalized du Bois-Reymond lemma.**

**Theorem 2.1** (du Bois-Reymond lemma). *Suppose $f,g\in L^1([a,b])$, $\delta\in L^\infty([a,b])$ and $\delta(s)>0$ for almost all $s\in[a,b]$. Set*

$$\Omega=\{\beta\in L^\infty([a,b]):\int_a^b\beta(s)\,ds=0,|\beta|\leqslant\delta,a.e.\}.$$

*If*

$$\int_a^b f(s)b_\beta(s)+g(s)\beta(s)\,ds=0,\quad \beta\in\Omega, \tag{2.1}$$

*where $b_\beta(s):=\int_a^s\beta(r)\,dr$ for $\beta\in\Omega$, then there exists a continuous representative $\tilde g$ of $g$ such that $\tilde g$ is absolutely continuous on $[a,b]$ and $\tilde g'(s)=f(s)$ for almost all $s\in[a,b]$.*

*Proof.* Without loss of generality, we suppose that $\delta\in L^\infty([a,b])$ and $\delta(s)>0$ for all $s\in[a,b]$.

Suppose first that $f\equiv0$. Let $\beta\in L^\infty([a,b])$, $\|\beta\|_\infty\leqslant1$, and $\int_a^b\beta\,ds=0$. Set $A^+=\{\beta\geqslant0\}$, $A^-=\{\beta<0\}$.

For any $n \geqslant 1$, by Lusin's theorem, there exists a compact set $E_n \subset [a,b]$ such that $|E_n| > (b-a) - 1/n$ and the restriction of $\delta$ on $E_n$ is continuous. Thus,

$$N_n := \min\{\delta(s) : s \in E_n\} > 0. \tag{2.2}$$

Set $A_n^{\pm} = A^{\pm} \cap E_n$. Consider the measure $\mu(E) = \int_E |\beta| \, ds$ which is absolutely continuous with respect to Lebesgue measure. We remark that $g \in L^1([a,b],\mu)$. Therefore, for any $\varepsilon > 0$ there exists $\sigma_\varepsilon > 0$ such that for any measurable subset $E \subset [a,b]$ we have that

$$\int_E |\beta| \, ds < \sigma_\varepsilon \quad \text{implies} \quad \int_E |g\beta| \, ds = \int_E |g| \, d\mu < \varepsilon/2. \tag{2.3}$$

Fix $\varepsilon > 0$ and let $\sigma = \sigma_\varepsilon$. Without loss of generality, we suppose that $\int_{A_n^+} |\beta| \, ds > \int_{A_n^-} |\beta| \, ds$[1]. The other possibility that $\int_{A_n^+} |\beta| \, ds < \int_{A_n^-} |\beta| \, ds$ can be dealt with in a similar way. Then, by (2.3), there exists $n_\varepsilon \in \mathbb{N}$ such that for all $n \geqslant n_\varepsilon$ we have that

$$0 < \int_{A_n^+} |\beta| \, ds - \int_{A_n^-} |\beta| \, ds < \sigma,$$

and

$$\int_{[a,b]\setminus E_n} |g| \, ds < \varepsilon/2. \tag{2.4}$$

Now, define $\psi(s) = \int_a^s |\beta| \cdot \mathbb{1}_{A_n^+} \, dr$. Then $\psi(b) > \int_{A_n^-} |\beta| \, ds$ and $\psi(a) = 0$. Taking $s_n = \sup\{s \in [a,b] : \psi(s) < \int_{A_n^-} |\beta| \, ds\}$, we have that

$$\int_a^{s_n} |\beta(r)| \cdot \mathbb{1}_{A_n^+}(r) \, dr = \int_{A_n^-} |\beta(s)| \, ds,$$

and, for $n \geqslant n_\varepsilon$ we have that

$$\int_{s_n}^b |\beta| \cdot \mathbb{1}_{A_n^+} \, dr = \int_{A_n^+} |\beta| \, ds - \int_{A_n^-} |\beta| \, ds < \sigma. \tag{2.5}$$

Define

$$\beta_n(s) = \begin{cases} \beta(s), & s \in J_n := (A_n^+ \cap [a, s_n]) \cup A_n^-; \\ 0, & \text{otherwise,} \end{cases}$$

where $\cup$ stands for the union of two disjoint set. Then $N_n\beta_n \in \Omega$ with $N_n$ defined in (2.2). Next, suppose $f \equiv 0$. Then, in light of (2.1), we have that

$$\int_a^b g\beta_n \, ds = 0. \tag{2.6}$$

Notice that $E_n = J_n \cup (A_n^+ \cap (s_n, b])$, or equivalently, $[a,b] \setminus J_n = ([a,b] \setminus E_n) \cup (A_n^+ \cap (s_n, b])$. Therefore, invoking (2.6), (2.4), (2.5) and (2.3) and recalling that the integral of $g\beta_n$ vanishes by our assumption, we conclude that for $n \geqslant n_\varepsilon$

$$\begin{aligned} \left| \int_a^b g\beta \, ds \right| &\leqslant \left| \int_{J_n} g\beta \, ds \right| + \left| \int_{[a,b]\setminus J_n} g\beta \, ds \right| \leqslant \left| \int_a^b g\beta_n \, ds \right| + \int_{[a,b]\setminus J_n} |g\beta| \, ds \\ &\leqslant \int_{[a,b]\setminus E_n} |g\beta| \, ds + \int_{s_n}^b |g\beta| \cdot \mathbb{1}_{A_n^+} \, ds \\ &< \varepsilon. \end{aligned}$$

[1] If the two integrals are equal, we go directly to $\beta_n$ below.

Since $\varepsilon$ is arbitrary we conclude that (2.1), in the case of $f \equiv 0$, holds for any $\beta \in L^\infty([a,b])$ such that $\|\beta\|_\infty \leqslant 1$ and $\int_a^b \beta\, ds = 0$. Now, our conclusion is a consequence of the standard du Bois-Reymond lemma (see, for instance, [19, Lemma 6.1.1]).

Finally, to treat the case of $f \not\equiv 0$, let $F(s) = \int_a^s f(r)\, dr$. Then $F' = f$ almost everywhere on $[a,b]$. Notice that $b_\beta(a) = b_\beta(b) = 0$ for any $\beta \in \Omega$. Then, by (2.6), for any $\beta \in \Omega$ we obtain that

$$\begin{aligned} 0 &= \int_a^b f(s)b_\beta(s) + g(s)\beta(s)\, ds \\ &= \int_a^b \frac{d}{ds}(F(s)b_\beta(s))\, ds + \int_a^b (g(s) - F(s))\beta(s)\, ds \\ &= \int_a^b (g(s) - F(s))\beta(s)\, ds. \end{aligned}$$

By the first step, we conclude $g - F$ is a.e. equal to some constant $c_0$. So, $\tilde{g} = F + c_0$. □

## 2.2. **Erdmann condition.**

**Theorem 2.2** (Erdmann condition)**.** *Suppose* (L1)-(L4) *are satisfied. Let* $\xi \in \Gamma^{a,b}_{x,y}$ *be a minimizer of* (1.4) *with* $u_\xi$ *determined by* (1.3)*. Set*

$$\int_a^s L_u dr = \int_a^s L_u(r, \xi(r), \dot{\xi}(r), u_\xi(r))dr$$

*and define*

$$\begin{aligned} E(s) :=& e^{-\int_a^s L_u dr} \cdot \Big\{ L_v(s, \xi(s), \dot{\xi}(s), u_\xi(s)) \cdot \dot{\xi}(s) - L(s, \xi(s), \dot{\xi}(s), u_\xi(s)) \Big\} \\ =& e^{-\int_a^s L_u dr} \cdot E_0(s) \end{aligned}$$

*for almost all* $s \in [a,b]$*. Then* $E$ *has a continuous representation* $\bar{E}$ *such that* $\bar{E}$ *is absolutely continuous on* $[a,b]$ *and*

$$\frac{d}{ds}\,\bar{E}(s) = -e^{-\int_a^s L_u dr} L_t(s) \tag{2.7}$$

*for almost all* $s \in [a,b]$*, where* $L_t(s) = L_t(s, \xi(s), \dot{\xi}(s), u_\xi(s))$*.*

*Proof.* We divide the proof into several steps. Without loss of generality, we suppose the time interval is $[0,t]$ with $t = b - a$.

**Step I: Reparameterization.** We follow the approach from [21]. For any measurable function $\alpha : [0,t] \to [1/2, 3/2]$ satisfying $\int_0^t \alpha(s)\, ds = t$ (the set of all such functions $\alpha$ is denoted by $\Omega$), we define

$$\tau(s) = \int_0^s \alpha(r)\, dr, \quad s \in [0,t].$$

Note that $\tau : [0,t] \to [0,t]$ is a bi-Lipschitz map and its inverse $s(\tau)$ satisfies

$$s'(\tau) = \frac{1}{\alpha(s(\tau))}, \quad a.e.\ \tau \in [0,t].$$

Now, given $\xi \in \Gamma^{0,t}_{x,y}$ as above and $\alpha \in \Omega$, define the reparameterization $\eta$ of $\xi$ by $\eta(\tau) = \xi(s(\tau))$. It follows that $\dot{\eta}(\tau) = \dot{\xi}(s(\tau))/\alpha(s(\tau))$. Let $u_\eta$ be the unique solution of (1.3) with initial condition $u_\eta(0)=u$. Then we have that

$$\begin{aligned}J(\xi)\leqslant\ J(\eta)&=\int_0^t L(\tau,\eta(\tau),\dot\eta(\tau),u_\eta(\tau))\,d\tau\\&=\int_0^t L(\tau(s),\xi(s),\dot\xi(s)/\alpha(s),u_{\xi,\alpha}(s))\alpha(s)\,ds\end{aligned}$$

where $u_{\xi,\alpha}$ solves

$$\dot u_{\xi,\alpha}(s)=L(\tau(s),\xi(s),u_{\xi,\alpha}(s),\dot\xi(s)/\alpha(s))\alpha(s),\quad u_{\xi,\alpha}(0)=u. \tag{2.8}$$

By a direct calculation, for all $\alpha\in\Omega$ and almost all $s\in[0,t]$, we obtain

$$\begin{aligned}\dot u_{\xi,\alpha}-\dot u_\xi&=L(\tau,\xi,\dot\xi/\alpha,u_{\xi,\alpha})\alpha-L(s,\xi,\dot\xi,u_\xi)\\&=L(\tau,\xi,\dot\xi/\alpha,u_{\xi,\alpha})\alpha-L(\tau,\xi,\dot\xi/\alpha,u_\xi)\alpha\\&\quad+L(\tau,\xi,\dot\xi/\alpha,u_\xi)\alpha-L(s,\xi,\dot\xi,u_\xi)\\&=\widehat{L_u^\alpha}\,(u_{\xi,\alpha}-u_\xi)+(L(\tau,\xi,\dot\xi/\alpha,u_\xi)\alpha-L(s,\xi,\dot\xi,u_\xi))\end{aligned}$$

and $u_{\xi,\alpha}(0)-u_\xi(0)=0$, where

$$\widehat{L_u^\alpha}(s)=\int_0^1 L_u\big(\tau(s),\xi(s),\dot\xi(s)/\alpha(s),u_\xi(s)+\lambda(u_{\xi,\alpha}(s)-u_\xi(s))\big)\alpha(s)\,d\lambda.$$

By solving the Carathéodory equation above, we conclude that

$$u_{\xi,\alpha}(s)-u_\xi(s)=\int_0^s e^{\int_\sigma^s\widehat{L_u^\alpha}\,dr}(L(\tau,\xi,\dot\xi/\alpha,u_\xi)\alpha-L(\sigma,\xi,\dot\xi,u_\xi))\,d\sigma \tag{2.9}$$

and $u_{\xi,\alpha}(t)-u_\xi(t)\geqslant 0$ for all $\alpha\in\Omega$.

**Step II: Summability after reparameterization.** For $\alpha\in[1/2,3/2]$ we define

$$\Phi_1(s,\alpha):=L(s,\xi(s),\dot\xi(s)/\alpha,u_\xi(s))\alpha-L(s,\xi(s),\dot\xi(s),u_\xi(s)).$$

For almost all $s$, by continuity, there exists $\delta_1(s)\in(0,1/2]$ such that

$$-1\leqslant\Phi_1(s,\alpha)-\Phi_1(s,1)\leqslant 1,\quad\forall\alpha\in[1-\delta_1(s),1+\delta_1(s)].$$

We define a set-valued map $G:[0,t]\rightrightarrows\mathbb{R}$ by

$$[0,t]\ni s\mapsto G(s)=\{\delta>0:\Phi_1(s,[1-\delta,1+\delta])\subset\Phi_1(s,1)+[-1,1]\},$$

and for each $k\in\mathbb{N}$ a set-valued map $G_k:[0,t]\rightrightarrows\mathbb{R}$ by

$$\begin{aligned}\mathrm{dom}\,(G_k)\ni s\mapsto&\,G_k(s)\\&=\{\delta\geqslant 1/k:\Phi_1(s,[1-\delta,1+\delta])\subset\Phi_1(s,1)+[-1,1]\}.\end{aligned}$$

By a standard measurable selection theorem (see, for instance, [22]), for each $k$, there exists a measurable selection $g_k:\mathrm{dom}\,(G_k)\to\mathbb{R}$ such that $g_k(s)\in G_k(s)$ for all $s\in[0,t]\cap\mathrm{dom}\,(G_k)$. Notice that we can assume that the sequence $\{g_k\}$ is nondecreasing and converges to a measurable selection $g$ of $G$ as $k\to\infty$. Thus, we can assume $\delta(\cdot)$ is measurable and $\delta(s)>0$ for almost all $s\in[0,t]$. We conclude that, if $\alpha\in\Omega$ satisfies $|\alpha(s)-1|\leqslant\delta(s)$ almost all $s\in[0,t]$, then

$$L(s,\xi(s),\dot\xi(s)/\alpha(s),u_\xi(s))\alpha(s)\in L^1([0,t]). \tag{2.10}$$

Set

$$\Omega_0=\{\alpha\in\Omega:|\alpha(s)-1|<\delta(s)\text{ a.e }s\in[0,t]\}$$

For any $M>0$ we define $E_M=\{s\in[0,t]:\dot\xi(s)$ exists and $|\dot\xi(s)|\leqslant M\}$. We choose $M$ such that $|E_M|>0$ and $|E_M^c|>0$ and set $\beta_M=|E_M|\cdot\mathbb{1}_{E_M^c}-|E_M^c|\cdot\mathbb{1}_{E_M}$. Notice that $\alpha=1+\varepsilon\beta_M$ satisfies the summability condition (2.10) for small $\varepsilon>0$ even if we cannot ensure $1+\varepsilon\beta_M\in\Omega_0$.

Fix $s\in[0,t]$ such that $\dot\xi$ exists and $\alpha\in\Omega_0$. Given $a\in\mathbb{R}$ we define $f(\lambda)=L(s+\lambda a,\xi(s),\dot\xi(s)/\alpha(s),u_\xi(s))$ for $\lambda\in[0,t]$. Without loss of generality, we suppose $f(\lambda)\geqslant 0$ for all $\lambda\in[0,t]$ by Proposition A.4 and condition (L2). Invoking condition (L4) we obtain that for all $\lambda\in[0,1]$

$$\begin{aligned}f'(\lambda)=&L_t(s+\lambda a,\xi(s),\dot\xi(s)/\alpha(s),u_\xi(s))\cdot a\\ \leqslant&[C_1+C_2L(s+\lambda a,\xi(s),\dot\xi(s)/\alpha(s),u_\xi(s))]\cdot a=[C_1+C_2f(\lambda)]\cdot a\\ \leqslant&C_1|a|+C_2|a|f(\lambda).\end{aligned}$$

Applying Gronwall's inequality we have that for all $\lambda\in[0,1]$

$$f(\lambda)\leqslant e^{C_2|a|\lambda}f(1)+C_1|a|\int_0^\lambda e^{C_2|a|(\lambda-s)}\,ds<e^{C_2|a|}(f(1)+C_1|a|).$$

It follows that

$$L(s+\varepsilon(\tau(s)-s),\xi(s),\dot\xi(s)/\alpha(s),u_\xi(s))\text{ is bounded by a function in }L^1([0,t])\tag{2.11}$$

for any $\varepsilon\in[0,1]$ provided $\alpha\in\Omega_0$ or $\alpha=1+\varepsilon\beta_M$.

**Step III: A necessary condition.** Fix $0\neq\beta\in L^\infty([0,t])$ such that $1+\beta\in\Omega_0$. For any $\varepsilon\in\mathbb{R}$ such that $|\varepsilon|\leqslant 1$ we have that $1+\varepsilon\beta\in\Omega_0\subset\Omega$. Let $\gamma(s)=\int_0^s\beta(r)\,dr$. Define the functional $\Lambda:\Omega\to\mathbb{R}$ by

$$\Lambda(\alpha)=u_{\xi,\alpha}(t)$$

with $u_{\xi,\alpha}$ defined in (2.8). Since $\Lambda(1+\varepsilon\beta)\geqslant\Lambda(1)$ for $|\varepsilon|\leqslant 1$, we have that $\frac{d}{d\varepsilon}\Lambda(1+\varepsilon\beta)|_{\varepsilon=0}=0$ if the derivative exists. Thus, for $\varepsilon>0$, by (2.9),

$$0\leqslant\frac{\Lambda(1+\varepsilon\beta)-\Lambda(1)}{\varepsilon}=\int_0^t e^{\int_s^t\widehat{L_u^\varepsilon}\,dr}\lambda_\varepsilon(s)\,ds,\tag{2.12}$$

where $\widehat{L_u^\varepsilon}=\widehat{L_u^{1+\varepsilon\beta}}$ and

$$\lambda_\varepsilon(s):=\frac{L(s+\varepsilon\gamma,\xi,\dot\xi/(1+\varepsilon\beta),u_\xi)(1+\varepsilon\beta)-L(s,\xi,\dot\xi,u_\xi)}{\varepsilon}.\tag{2.13}$$

Set

$$l_\varepsilon(s):=L_v(s,\xi,\dot\xi/(1+\varepsilon\beta),u_\xi)\cdot\dot\xi/(1+\varepsilon\beta)-L(s,\xi,\dot\xi/(1+\varepsilon\beta),u_\xi).$$

For convenience we take out the variable $s$ on right side of the inequalities above. We claim that

$$0=\frac{d}{d\varepsilon}\Lambda(1+\varepsilon\beta)|_{\varepsilon=0}=\int_0^t e^{\int_s^t L_u\,dr}\left\{l_0\cdot\beta-L_t(s,\xi,\dot\xi,u_\xi)\cdot\gamma\right\}\,ds.\tag{2.14}$$

**Step IV: On the summability.** By convexity we have that

$$\begin{aligned}&L(s,\xi,\dot\xi/(1+\varepsilon\beta),u_\xi)-L(s,\xi,\dot\xi,u_\xi)\\ \leqslant&-L_v(s,\xi,\dot\xi/(1+\varepsilon\beta),u_\xi)\cdot\{\dot\xi-\dot\xi/(1+\varepsilon\beta)\}\\ =&-\varepsilon\beta L_v(s,\xi,\dot\xi/(1+\varepsilon\beta),u_\xi)\cdot\dot\xi/(1+\varepsilon\beta).\end{aligned}$$

It follows that

$$
\begin{aligned}
\lambda_\varepsilon \leqslant & -\beta\{L_v(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)\cdot\dot{\xi}/(1+\varepsilon\beta)-L(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)\}\\
& +\beta\{L(s+\varepsilon\gamma,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)-L(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)\}\\
& +\frac{1}{\varepsilon}(L(s+\varepsilon\gamma,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)-L(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi))\\
= & -\beta\cdot l_\varepsilon+\beta\cdot b_\varepsilon+\frac{b_\varepsilon}{\varepsilon},
\end{aligned}
\tag{2.15}
$$

where

$$
b_\varepsilon = L(s+\varepsilon\gamma,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)-L(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi).
$$

In order to validate the integrand on the right side of (2.12), we need to focus on the summability of $\lambda_\varepsilon$. We observe that

$$
L(s+\varepsilon\gamma,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)-L(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)=\int_0^1\widehat{L_t}(\lambda,s)\,d\lambda\cdot\varepsilon\gamma,
$$

where for $\lambda\in[0,1]$ we denote

$$
\widehat{L_t}(\lambda,s):=L_t(s+\lambda\varepsilon\gamma,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)
$$

Due to condition (L4), we have that

$$
\begin{aligned}
&\frac{1}{\varepsilon}\left|L(\tau,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)-L(s,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)\right|\\
\leqslant & |\gamma|\cdot\int_0^1 C_1+C_2L(s+\lambda\varepsilon\gamma,\xi,\dot{\xi}/(1+\varepsilon\beta),u_\xi)\,d\lambda.
\end{aligned}
$$

Thus $b_\varepsilon(s)/\varepsilon$ is bounded by an $L^1$-function by (2.11).

In view of Proposition A.2 (a), we have that

$$
l_\varepsilon\geqslant -L(s,\xi,0,u_\xi)\geqslant -\overline{\theta}_0(0)-c_1-KF(t,|y-x|/t).
$$

For any $\beta\in\Omega_0$ and $\varepsilon\in[0,1]$, we rewrite $\lambda_\varepsilon(s)$, $l_\varepsilon(s)$ and $b_\varepsilon(s)$ as $\lambda^\beta_\varepsilon(s)$, $l^\beta_\varepsilon(s)$ and $b^\beta_\varepsilon(s)$ respectively.

Set $\beta^+=\beta\cdot\mathbb{1}_{\{\beta\geqslant 0\}}$ and $\beta^-=-\beta\cdot\mathbb{1}_{\{\beta<0\}}$, then

$$
\beta=\beta^+-\beta^-,\quad\text{and}\quad \beta^\pm\geqslant 0.
$$

By (2.15) we have that

$$
\lambda^\beta_\varepsilon(s)+\beta^+(s)l^\beta_\varepsilon(s)-\beta(s)b^\beta_\varepsilon(s)-\frac{b^\beta_\varepsilon(s)}{\varepsilon}\leqslant\beta^-(s)l^\beta_\varepsilon(s).
$$

Now, observe that $\beta^+(s)l^\beta_\varepsilon(s)=\beta^+(s)l^{\beta^+}_\varepsilon(s)$ and $\beta^-(s)l^\beta_\varepsilon(s)=\beta^-(s)l^{\beta^-}_{-\varepsilon}(s)$. Then the inequalities above can recast as follows

$$
\lambda^\beta_\varepsilon(s)+\beta^+(s)l^{\beta^+}_\varepsilon(s)-\beta(s)b^\beta_\varepsilon(s)-\frac{b^\beta_\varepsilon(s)}{\varepsilon}\leqslant\beta^-(s)l^{\beta^-}_{-\varepsilon}(s).\tag{2.16}
$$

Lemma A.2 (a) ensures that $\varepsilon\mapsto l^{\beta^-}_\varepsilon$ is decreasing on $[-1,1]$ and we conclude that

$$
\beta^-l^{\beta^-}_{-\varepsilon}\leqslant\beta^-l^{\beta^-}_{-1}\quad\forall\varepsilon\in(0,1).\tag{2.17}
$$

By Lemma A.2 (b), we obtain

$$
\begin{aligned}
\beta^-l^{\beta^-}_{-\varepsilon}=&\beta^-\{L_v(s,\xi,\dot{\xi}/(1-\varepsilon\beta^-),u_\xi)\cdot\dot{\xi}/(1-\varepsilon\beta^-)-L(s,\xi,\dot{\xi}/(1-\varepsilon\beta^-),u_\xi)\}\\
\leqslant&(\kappa^{\beta^-}_\varepsilon)^{-1}L(s,\xi,\dot{\xi}/(1-\beta^-),u_\xi)-((\kappa^{\beta^-}_\varepsilon)^{-1}+\beta^-)L(s,\xi,\dot{\xi}/(1-\varepsilon\beta^-),u_\xi)
\end{aligned}
$$

where $(\kappa_\varepsilon^{\beta^-})^{-1}=\frac{1-\beta^-}{1-\varepsilon}$. In view of (2.11), (2.16) and the fact that $(\kappa_\varepsilon^{\beta^-})^{-1}$ is bounded, we conclude that $\beta^- l_{-\varepsilon}^{\beta^-}\in L^1([0,t])$ for all $\varepsilon\in(0,1]$ uniformly.

**Step IV: Erdmann condition.** We rewrite $L_t(s)=L_t(s,\xi(s),\dot\xi(s),u_\xi(s))$. Recalling that for almost all $s\in[0,t]$ we have that

$$\lim_{\varepsilon\to0^+} b_\varepsilon^\beta(s)=0,\quad \lim_{\varepsilon\to0^+}\frac{b_\varepsilon^\beta(s)}{\varepsilon}=L_t(s)\cdot\gamma(s).$$

Thus, integrating (2.16), by Lebesgue's theorem we obtain

$$\int_0^t e^{\int_s^t L_u dr}\{l_0(s)\beta^+(s)-L_t(s)\gamma(s)\}\,ds\leqslant\int_0^t e^{\int_s^t L_u dr} l_0(s)\beta^-(s)ds.$$

Therefore, $\int_0^t e^{\int_s^t L_u dr}\{l_0\cdot\beta-L_t\cdot\gamma\}\,ds\leqslant0$ and (2.14) follows since $\beta\in\Omega_0$ is arbitrary.

Now, observe that the primitive $\mu(s):=\int_0^s\beta(r)dr$ gives a one-to-one correspondence between $\Omega_0$ and the set

$$\Omega_1=\{\mu:[0,t]\to\mathbb{R}:\mu \text{ is Lipschitz continuous with } \mu(0)=\mu(t)=0,\mu'\in\Omega_0\}.$$

Thus, (2.14) can be recast as follows

$$0=-e^{\int_0^t L_u dr}\int_0^t E(s)\mu'(s)-e^{-\int_0^s L_u dr}L_t(s)\mu(s)\,ds\quad\forall\mu\in\Omega_1.$$

So, (2.7) follows by the generalized du Bois-Reymond lemma [2] (Theorem 2.1). □

## 2.3. **Herglotz equation.**

**Theorem 2.3** (Herglotz equation)**.** *Suppose conditions* (L1)-(L3) *and* (L4') *are satisfied. Let* $\xi\in\Gamma_{x,y}^{a,b}$ *be a minimizer of* (1.4) *with* $u_\xi$ *determined by* (1.3)*. Then, the function* $s\mapsto L_x(s,\xi(s),\dot\xi(s),u_\xi(s))$ *is absolutely continuous on* $[a,b]$ *and* $(\xi,u_\xi)$ *satisfies the Carathéodory equation* (1.3) *and the Herglotz equation*

$$\begin{aligned}&\frac{d}{ds}L_v(s,\xi(s),\dot\xi(s),u_\xi(s))\\&=L_x(s,\xi(s),\dot\xi(s),u_\xi(s))+L_u(s,\xi(s),\dot\xi(s),u_\xi(s))L_v(s,\xi(s),\dot\xi(s),u_\xi(s))\end{aligned}\tag{2.18}$$

*for almost all* $s\in[a,b]$.

*Proof.* Let $\xi\in\Gamma_{x,y}^{a,b}$ be a minimizer of (1.4) where $u_\xi$ is determined uniquely by (1.3). For any $\varepsilon\in\mathbb{R}$ and any Lipschitz function $\eta\in\Gamma_{0,0}^{a,b}$, we set $\xi_\varepsilon(s)=\xi(s)+\varepsilon\eta(s)$. Let $u_{\xi_\varepsilon}$ be the associated unique solution of (1.3) with respect to $\xi_\varepsilon$, i.e., $u_{\xi_\varepsilon}$ satisfies

$$\begin{cases}\dot u_{\xi_\varepsilon}(s)=L(s,\xi_\varepsilon(s),\dot\xi_\varepsilon(s),u_{\xi_\varepsilon}(s)), & a.e.\ s\in[a,b],\\ u_{\xi_\varepsilon}(0)=u.\end{cases}\tag{2.19}$$

It is clear that $\xi_\varepsilon\in\Gamma_{x,y}^{a,b}$ and $J(\xi)\leqslant J(\xi_\varepsilon)$. Combining (2.19) and (1.3) we have that

$$\begin{aligned}\dot u_{\xi_\varepsilon}-\dot u_\xi&=L(s,\xi_\varepsilon,\dot\xi_\varepsilon,u_{\xi_\varepsilon})-L(s,\xi,\dot\xi,u_\xi)\\&=\{L(s,\xi_\varepsilon,\dot\xi_\varepsilon,u_{\xi_\varepsilon})-L(s,\xi_\varepsilon,\dot\xi_\varepsilon,u_\xi)\}+\{L(s,\xi_\varepsilon,\dot\xi_\varepsilon,u_\xi)-L(s,\xi,\dot\xi,u_\xi)\}\\&=\widehat{L_u^\varepsilon}(u_{\xi_\varepsilon}-u_\xi)+\{L(s,\xi_\varepsilon,\dot\xi_\varepsilon,u_\xi)-L(s,\xi,\dot\xi,u_\xi)\},\end{aligned}$$

[2]By (2.11), the previous steps of the proof can also be applied to $\beta_M$. This shows that (2.14) holds for $\beta_M$ and this leads to the summability of $l_0$ as well as $E$. This allows us to use Theorem 2.1.

where

$$\widehat{L^{\varepsilon}_{u}}(s)=\int_0^1 L_u(s,\xi_\varepsilon(s),\dot{\xi}_\varepsilon(s),u_\xi(s)+\lambda(u_{\xi_\varepsilon}(s)-u_\xi(s)))\,d\lambda.$$

It follows that

$$u_{\xi_\varepsilon}(s)-u_\xi(s)=\int_a^s e^{\int_\sigma^s \widehat{L^{\varepsilon}_{u}}\,dr}(L(s,\xi_\varepsilon,\dot{\xi}_\varepsilon,u_\xi)-L(s,\xi,\dot{\xi},u_\xi))\,d\sigma.$$

Recalling that $J(\xi_\varepsilon)=u_\xi(t)$, we obtain

$$0\leqslant\frac{J(\xi_\varepsilon)-J(\xi)}{\varepsilon}=\int_a^b e^{\int_s^b \widehat{L^{\varepsilon}_{u}}\,dr}\cdot\frac{L(s,\xi_\varepsilon,\dot{\xi}_\varepsilon,u_\xi)-L(s,\xi,\dot{\xi},u_\xi)}{\varepsilon}\,ds. \tag{2.20}$$

Now, similarly to Step II of the proof of Theorem 2.2, by using the measurable selection theorem, there exists $\delta\in L^\infty([a,b])$, with $\delta>0$ a.e., such that, if $|\eta(s)|\leqslant\delta(s)$ for almost all $s\in[a,b]$, then $L(s,\xi_\varepsilon,\dot{\xi}_\varepsilon,u_\xi)$ is bounded by an $L^1$-function uniformly for $|\varepsilon|\leqslant 1$. Invoking condition (L4'), we conclude that $L_x(s,\xi_\varepsilon,\dot{\xi},u_\xi)$ is also bounded by an $L^1$-function uniformly for $|\varepsilon|\leqslant 1$. By convexity we have that

$$L(s,\xi,\dot{\xi},u_\xi)-L(s,\xi,\dot{\xi}_{-1},u_\xi)\leqslant L_v(s,\xi,\dot{\xi},u_\xi)\cdot\dot{\eta}\leqslant L(s,\xi,\dot{\xi}_1,u_\xi)-L(s,\xi,\dot{\xi},u_\xi).$$

It follows that $L_v(s,\xi,\dot{\xi},u_\xi)\cdot\dot{\eta}\in L^1([a,b])$. Now, we can assume that $L_v(s,\xi,\dot{\xi}_\varepsilon,u_\xi)\cdot\dot{\eta}$ is bounded by an $L^1$-function for all $|\varepsilon|\leqslant 1$.

Fix $\eta\in\Gamma^{a,b}_{0,0}$ such that $|\eta(s)|\leqslant\delta(s)$ for almost all $s\in[0,t]$. We claim that

$$\frac{d}{d\varepsilon}J(\xi_\varepsilon)=0=\int_a^b e^{\int_s^b L_u\,dr}\cdot\{L_x\cdot\eta+L_v\cdot\dot{\eta}\}\,ds. \tag{2.21}$$

By convexity, we have that

$$L_v(s,\xi,\dot{\xi},u_\xi)\cdot\dot{\eta}\leqslant\frac{L(s,\xi,\dot{\xi}_\varepsilon,u_\xi)-L(s,\xi,\dot{\xi},u_\xi)}{\varepsilon}\leqslant L_v(s,\xi,\dot{\xi}_\varepsilon,u_\xi)\cdot\dot{\eta}.$$

Moreover,

$$\left|\frac{L(s,\xi_\varepsilon,\dot{\xi}_\varepsilon,u_\xi)-L(s,\xi,\dot{\xi}_\varepsilon,u_\xi)}{\varepsilon}\right|\leqslant|\eta|\int_0^1|L_x(s,\xi_\varepsilon+\lambda(\xi_\varepsilon-\xi),\dot{\xi}_\varepsilon,u_\xi)|\,d\lambda$$

Taking the limit in (2.20) as $\varepsilon\to 0^+$, then (2.21) follows by Lebesgue's theorem. Thus, (2.18) follows by Theorem 2.1 provided $L_v(s,\xi(s),\dot{\xi}(s),u_\xi(s))\in L^1([a,b])$ which is guaranteed by condition (L4'). □

*Remark* 2.4. It is also useful to rewrite the Herglotz equation is the form

$$\frac{d}{ds}e^{-\int_a^s L_u(r)\,dr}L_v(s,\xi(s),\dot{\xi}(s),u_\xi(s))=e^{-\int_a^s L_u(r)\,dr}L_x(\xi(s),\dot{\xi}(s),u_\xi(s)), \tag{2.22}$$

where $L_u(s)=L_u(s,\xi(s),\dot{\xi}(s),u_\xi(s))$.

2.4. **Lipschitz estimates.** In this section, we will prove the Lipschitz estimates for the minimizer $\xi$ of (1.4).

**Theorem 2.5.** *Suppose conditions* (L1)-(L3) *are satisfied together with either* (L4) *or* (L4'). *Let $u\in\mathbb{R}$ and $R>0$ be fixed. Then there exists a continuous function $F=F_{u,R}:[0,+\infty)\times[0,+\infty)\to[0,+\infty)$, with $F(t,r)$ nondecreasing in both variables* *and superlinear with respect to $r$, such that for any given $b > a$ and $x, y \in \mathbb{R}^n$, with $|x - y| \leqslant R$, every minimizer $\xi \in \Gamma^{a,b}_{x,y}$ for* (1.4) *satisfies*

$$\operatorname*{ess\,sup}_{s\in[a,b]} |\dot{\xi}(s)| \leqslant F(b-a, R/(b-a)).$$

*Proof.* We consider two cases, one for each of the different assumptions of the theorem.

**Case I**: We assume conditions (L1)-(L3) together with (L4).

Let $\xi \in \Gamma^{a,b}_{x,y}$ be a minimizer of (1.4), for $\alpha > 0$. Set

$$l_\xi(s,\alpha) = \alpha \cdot L(s, \xi(s), \dot{\xi}(s)/\alpha, u_\xi(s))$$

and recall $E_0 = L_v(s,\xi,\dot{\xi},u_\xi)\cdot\dot{\xi} - L(s,\xi,\dot{\xi},u_\xi)$. Simple computations show that $l_\xi(s,\cdot)$ is convex and

$$\frac{d}{d\alpha}\Big|_{\alpha=1} l_\xi(s,\alpha) = -E_0(s).$$

Choosing $s_0 \in [a,b]$ such that $|\dot{\xi}(s_0)| = \operatorname{ess\,inf}_{s\in[a,b]} |\dot{\xi}(s)|$ by convexity, we have that

$$-E_0(s_0) \geqslant \sup_{\alpha<1} \frac{l_\xi(s_0,1) - l_\xi(s_0,\alpha)}{1-\alpha}$$

Recall that $|u_\xi|$ is bounded by $F_1(b-a, R/(b-a))$ and $\operatorname{ess\,inf}_{s\in[a,b]} |\dot{\xi}(s)|$ is bounded by $F_2(b-a, R/(b-a))$ by Proposition A.4. For convenience, we drop the variables in the functions $F_1$ and $F_2$, and also $F_i$ in the following text.

Taking $\alpha = \frac{1}{2}$, by (L2)-(L3) we conclude that

$$\begin{aligned}
-E_0(s_0) \geqslant& 2(l_\xi(s_0,1) - l_\xi(s_0,1/2)) = 2(L(s_0,\xi(s_0),\dot{\xi}(s_0),u_\xi(s_0)) - l_\xi(s_0,1/2))\\
\geqslant& -2c_0 - 2KF_1 - L(s_0,\xi(s_0),2\dot{\xi}(s_0),u_\xi(s_0))\\
\geqslant& -2c_0 - 3KF_1 - L(s_0,\xi(s_0),2\dot{\xi}(s_0),0)\\
\geqslant& -2c_0 - 3KF_1 - \overline{\theta}_0(2|\dot{\xi}(s_0)|) - c_1\\
\geqslant& -2c_0 - 3KF_1 - \overline{\theta}_0(2F_2) - c_1 := -F_3.
\end{aligned}$$

We rewrite $L_t(s) = L_t(s,\xi(s),\dot{\xi}(s),u_\xi(s))$ and $L_u(s) = L_u(s,\xi(s),\dot{\xi}(s),u_\xi(s))$. Then, by Erdmann's condition (2.7) we obtain that for almost all $s \in [a,b]$,

$$\begin{aligned}
E(s) =& E(s_0) - \int_{s_0}^{s} e^{-\int_a^\tau L_u dr} L_t(\tau)\,d\tau \leqslant e^{-\int_a^{s_0} L_u dr} E_0(s_0) + \int_a^b e^{-\int_a^\tau L_u dr} |L_t(\tau)|\,d\tau\\
\leqslant& e^{K(b-a)}F_3 + e^{K(b-a)} \int_a^b |L_t(s)|\,ds.
\end{aligned}$$

By (L4) we conclude that

$$\begin{aligned}
E(s) \leqslant& e^{K(b-a)}F_3 + e^{K(b-a)} \int_a^b \big\{C_1 + C_2 L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\big\}\,ds\\
\leqslant& e^{K(b-a)}\big\{F_3 + C_1(b-a) + C_2F_4\big\} := F_5,
\end{aligned}$$

where $\int_a^b L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds$ is bounded by $F_4$ by Proposition A.4. Therefore, we have that, for almost all $s \in [a,b]$,

$$E_0(s) = e^{\int_a^s L_u d\tau} E(s) \leqslant e^{K(b-a)}F_5 := F_6. \tag{2.23}$$

Now, let $s$ be such that $\dot\xi(s)$ exists and (2.23) holds. By convexity, we have that

$$\begin{aligned}&L(s,\xi(s),\dot\xi(s)/(1+|\dot\xi(s)|),u_\xi(s))-L(s,\xi(s),\dot\xi(s),u_\xi(s))\\ \geqslant\,&((1+|\dot\xi(s)|)^{-1}-1)\cdot\langle L_v(s,\xi(s),\dot\xi(s),u_\xi(s)),\dot\xi(s)\rangle\\ \geqslant\,&((1+|\dot\xi(s)|)^{-1}-1)\cdot(L(s,\xi(s),\dot\xi(s),u_\xi(s))+F_6).\end{aligned}$$

It follows that

$$\begin{aligned}&L(s,\xi(s),\dot\xi(s),u_\xi(s))\\ \leqslant\,&L(s,\xi(s),\dot\xi(s)/(1+|\dot\xi(s)|),u_\xi(s))(1+|\dot\xi(s)|)+F_6|\dot\xi(s)|.\end{aligned}$$

Let $C=\sup_{s\in[a,b],|v|\leqslant1}L(s,\xi(s),v,u_\xi(s))$ and observe that, by (L2) and Proposition A.4,

$$C\leqslant\sup_{s\in[a,b],|v|\leqslant1}\{L(s,\xi(s),v,0)+K|u_\xi(s)|\}\leqslant\overline{\theta}_0(1)+c_1+KF_1:=F_7.$$

It follows that

$$L(s,\xi(s),\dot\xi(s),u_\xi(s))\leqslant F_7+(F_6+F_7)|\dot\xi(s)|.$$

Therefore, invoking Proposition A.4, we obtain

$$\begin{aligned}&(F_6+F_7+1)|\dot\xi(s)|-(\theta_0^*(F_6+F_7+1)+c_0)\\ \leqslant\,&\theta_0(|\dot\xi(s)|)-c_0\leqslant L(s,\xi(s),\dot\xi(s),0)\leqslant L(s,\xi(s),\dot\xi(s),u_\xi(s))+K|u_\xi(s)|\\ \leqslant\,&F_7+(F_6+F_7)|\dot\xi(s)|+KF_1.\end{aligned}$$

This leads to

$$|\dot\xi(s)|\leqslant(\theta_0^*(F_6+F_7+1)+c_0)+F_7+KF_1:=F_8,$$

which completes the proof of Case I.

**Case II**: We suppose conditions (L1)-(L3) together with (L4') are satisfied.

This case is much easier than Case I. Again, we choose $s_0\in[0,t]$ such that

$$|\dot\xi(s_0)|=\operatorname*{ess\,inf}_{s\in[a,b]}|\dot\xi(s)|\leqslant F_1.$$

By Corollary A.4, $\xi(s)$ is contained in $B(x,(b-a)F_2)$ and $|u_\xi(s)|$ is bounded by $F_3$. Set

$$F_4=\max\{|L_v(s_0,y,v,r)|:|y-x|\leqslant(b-a)F_2,|v|\leqslant F_1,|r|\leqslant F_3\}.$$

By solving Herglotz' equation in the form (2.22) we have that, for any $s\in[a,b]$,

$$\begin{aligned}&e^{-\int_a^s L_u\,d\tau}L_v(s,\xi(s),\dot\xi(s),u_\xi(s))\\ =\,&e^{-\int_a^{s_0} L_u\,d\tau}L_v(s_0,\xi(s_0),\dot\xi(s_0),u_\xi(s_0))+\int_{s_0}^s e^{-\int_a^\tau L_u\,dr}L_x\,d\tau\end{aligned}$$

By condition (L4') we conclude that, for almost $s\in[a,b]$,

$$\begin{aligned}&|L_v(s,\xi(s),\dot\xi(s),u_\xi(s))|\\ \leqslant\,&e^{2K(b-a)}F_4+e^{2K(b-a)}\int_a^b|L_x(s,\xi(s),\dot\xi(s),u_\xi(s))|\,ds\\ \leqslant\,&e^{2K(b-a)}F_4+e^{2K(b-a)}\int_a^b\big\{C_1+C_2L(s,\xi(s),\dot\xi(s),u_\xi(s))\big\}\,ds\\ \leqslant\,&e^{2K(b-a)}\big\{F_4+C_1(b-a)+C_2F_5\big\}:=F_6.\end{aligned}\tag{2.24}$$

Now, let $H$ be the Hamiltonian associated with $L$. Set

$$F_7 = \max\{|H_p(s, y, p, r)| : |y - x| \leqslant (b - a)F_2, |p| \leqslant F_6, |r| \leqslant F_3\}.$$

Then, for any $s \in [a, b]$ such that $\dot{\xi}(s)$ exists and (2.24) is satisfied, we obtain that

$$|\dot{\xi}(s)| = |H_p(s, \xi(s), L_v(s, \xi(s), \dot{\xi}(s), u_\xi(s)), u_\xi(s))| \leqslant F_7.$$

This completes the proof of Case II. □

**Corollary 2.6.** *Theorem 2.3 holds under the assumptions* (L1)-(L4) *or* (L1)-(L3) and (L4'). *In particular, Herglotz equation* (2.18) *holds true.*

*Proof.* Due to Theorem 2.5, we have the uniform bound of $\dot{\xi}(s)$ for almost all $s \in [0, t]$. Along the proof of Theorem 2.2, there is no summability difficulty since the Lipschitz estimates, and Erdmann condition (2.7) can be obtained directly by Step V in the proof of Theorem 2.2. Now, the proof of the theorem is similar to but simpler than that of Theorem 2.3 because of our Lipschitz estimates. □

**Corollary 2.7.** *The minimal curve $\xi$ of* (1.4) *is of class $C^1$ as well as $u_\xi$.*

*Proof.* Let $N$ be the set of zero Lebesgue measure where $\dot{\xi}$ does not exist. For $\bar{t} \in [a, b]$, choose a sequence $\{t_k\} \in [a, b] \setminus N$ such that $t_k \to \bar{t}$. Then $\dot{\xi}(t_k) \to \bar{v}$ for some $\bar{v} \in \mathbb{R}^n$ (up to subsequences) and

$$\begin{aligned}
&L_v(\bar{t}, \xi(\bar{t}), \bar{v}, u_\xi(\bar{t})) - L_v(t_1, \xi(t_1), \dot{\xi}(t_1), u_\xi(t_1)) \\
=& \lim_{k\to\infty} L_v(t_k, \xi(t_k), \dot{\xi}(t_k), u_\xi(t_k)) - L_v(t_1, \xi(t_1), \dot{\xi}(t_1), u_\xi(t_1)) \\
=& \int_{t_1}^{\bar{t}} \{L_x(s, \xi(s), \dot{\xi}(s), u_\xi(s)) + L_u(s, \xi(s), \dot{\xi}(s), u_\xi(s)) L_v(s, \xi(s), \dot{\xi}(s), u_\xi(s))\} ds
\end{aligned}$$

by Herglotz equation (2.18). From the strict convexity of $L$ it follows that the map $v \mapsto L_v(s, \xi(s), v, u_\xi(s))$ is a diffeomorphism. This implies that $\bar{v}$ is uniquely determined, i.e.,

$$\lim_{[0,t]\setminus N \ni s \to \bar{t}} \dot{\xi}(s) = \bar{v}.$$

Now, by Lemma 6.2.6 in [19], $\dot{\xi}(\bar{t})$ exists and $\lim_{[0,t]\setminus N\ni s\to\bar{t}} \dot{\xi}(s) = \dot{\xi}(\bar{t})$. It follows that $\xi$ is of class $C^1$. In view of (1.3), $u_\xi$ is also of class $C^1$. □

The following improvement of the main results in this section is very similar to that in [17]. We omit the proof.

**Proposition 2.8.** *Suppose $L$ is of class $C^2$ and satisfies conditions* (L1)-(L3) *together with* (L4) *or* (L4'). *For any fixed $x, y \in \mathbb{R}^n$, $b > a$ and $u \in \mathbb{R}$, the functional $J$ defined in* (1.4) *admits a minimizer. Moreover,*

(a) *both $\xi$ and $u_\xi$ are of class $C^2$ and $\xi$ satisfies Herglotz' equation* (1.2) *for all $s \in [a, b]$ where $u_\xi$ is the unique solution of* (1.3)*;*
(b) *the dual arc $p$ defined by $p(s) = L_v(s, \xi(s), \dot{\xi}(s), u_\xi(s))$ is also of class $C^2$ and $(\xi, p, u_\xi)$ satisfies Lie equation* (1.6) *for all $s \in [a, b]$.*

## 3. Applications to Hamilton-Jacobi equations of the contact type

In this section, we want to explain the relations between Herglotz' variational principle and the Hamilton-Jacobi equation (HJ). Throughout this section, we suppose that $L$ satisfies condition (L1)-(L3), together with (L4) or (L4'). Therefore Proposition 2.8 holds.

3.1. **Fundamental solutions and Lax-Oleinik evolution.** Fix $x, y \in \mathbb{R}^n$, $t_2 > t_1$ and $u \in \mathbb{R}$. Let $\xi \in \mathcal{A}_1 := \Gamma^{t_1,t_2}_{x,y} \cap C^2([t_1,t_2],\mathbb{R}^n)$ and let $u_\xi$ be the unique $C^2$ solution of the ODE

$$\begin{cases} \dot{u}_\xi(s) = L(s,\xi(s),\dot{\xi}(s),u_\xi(s)), & s \in [t_1,t_2], \\ u_\xi(t_1) = u. \end{cases} \tag{3.1}$$

We define

$$h_L(t_1,t_2,x,y,u) := \inf_{\xi\in\mathcal{A}_1} \int_{t_1}^{t_2} L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds = \inf_{\xi\in\mathcal{A}_1} u_\xi(t_2) - u. \tag{3.2}$$

An associated variational problem of Herglotz' type is as follows:

$$\breve{h}_L(t_1,t_2,x,y,u) := \inf_{\xi} \int_{t_1}^{t_2} L(s,\xi(s),\dot{\xi}(s),w_\xi(s))\,ds \tag{3.3}$$

where the infimum is taken over all $\xi \in \mathcal{A}_1$ such that a terminal condition problem of Carathéodory equation

$$\begin{cases} \dot{w}_\xi(s) = L(s,\xi(s),\dot{\xi}(s),w_\xi(s)), & s \in [t_1,t_2], \\ w_\xi(t_2) = u, \end{cases} \tag{3.4}$$

admits a (unique) solution. Invoking Proposition 2.8, the infimum in the definition of $h_L(t_1,t_2,x,y,u)$ and $\breve{h}_L(t_1,t_2,x,y,u)$ can be achieved.

**Definition 3.1.** Fix $x, y \in \mathbb{R}^n$, $t_2 > t_1$ and $u \in \mathbb{R}$. We call the function $h_L(t_1,t_2,x,y,u)$ (resp. $\breve{h}_L(t_1,t_2,x,y,u)$) the *negative* (resp. *positive*) *type fundamental solution* for (HJ).

**Definition 3.2** ($t$-dependent case)**.** For any function $\phi : \mathbb{R}^n \to [-\infty,+\infty]$, we define

$$\begin{aligned} (\mathbf{T}^{t_2}_{t_1}\phi)(x) &= \inf_{y\in\mathbb{R}^n}\{\phi(y) + h_L(t_1,t_2,y,x,\phi(y))\}, \\ (\breve{\mathbf{T}}^{t_2}_{t_1}\phi)(x) &= \sup_{y\in\mathbb{R}^n}\{\phi(y) - \breve{h}_L(t_1,t_2,x,y,\phi(y))\}, \end{aligned} \qquad t_2 > t_1, x \in \mathbb{R}^n.$$

The operators $\mathbf{T}^{t_2}_{t_1}$ and $\breve{\mathbf{T}}^{t_2}_{t_1}$ are called the *negative* and *positive type Lax-Oleinik operators*, respectively, and $\mathbf{T}^{t_2}_{t_1}\phi$ and $\breve{\mathbf{T}}^{t_2}_{t_1}\phi$ are called the *negative* and *positive type Lax-Oleinik evolution of* $\phi$, respectively.

**Definition 3.3.** Let $(x,d)$ be a metric space. A function $\phi : X \to \mathbb{R}$ is called $(\kappa_1,\kappa_2)$-Lipschitz in the large if there exists $\kappa_1,\kappa_2 \geqslant 0$ such that

$$|\phi(y) - \phi(x)| \leqslant \kappa_1 + \kappa_2 d(x,y), \quad \forall x, y \in X.$$

*Example* 3.4. Given $\phi : X \to \mathbb{R}$. We have that

(i) If $X$ is compact, it is obvious that $\phi$ is $(\kappa_1,\kappa_2)$-Lipschitz in the large if and only if $\phi$ is bounded.
(ii) If $X = \mathbb{R}^n$ or any complete Riemannian manifold and $\phi$ is uniformly continuous, then for any $\varepsilon > 0$ there exists $K_\varepsilon > 0$ such that $\phi$ is $(\varepsilon, K_\varepsilon)$-Lipschitz in the large (see Proposition A.3).
(iii) If $\phi$ is Lipschitz with constant $\mathrm{Lip}\,(\phi)$, then $\phi$ is $(0,\mathrm{Lip}\,(\phi))$-Lipschitz in the large.

*Remark* 3.5. We have some remarks on the operators $\mathbf{T}^{t_2}_{t_1}\phi$ and $\breve{\mathbf{T}}^{t_2}_{t_1}\phi$.

– Notice that there is no extra assumption on the function $\phi$ in Definition 3.2. But, to ensure that $\mathbf{T}^{t_2}_{t_1}\phi$ and $\breve{\mathbf{T}}^{t_2}_{t_1}\phi$ are finite-valued and the infimum and supremum in Definition 3.2 can be achieved, we need more conditions.

– In [6], the author pointed out that if $\phi$ is *continuous and Lipschitz in the large*, then $u(t,x)=(\mathbf{T}_0^t\phi)(x)$ is finite-valued for any classical time-dependent Lagrangian $L(t,x,v)$. For more informations on functions that are Lipschitz in the large and applications to Lax-Oleinik evolution in classical case, see [29, 15].
– Using an idea from the proof of Lemma 3.1 in [13] (see also [57] when the Lagrangian has the form $L(x,v,r)$), we can show that the infimum and supremum in Definition 3.2 can be achieved if $\phi$ is lower and upper semi-continuous respectively, and $(\kappa_1,\kappa_2)$-Lipschitz in the large. See Lemma 3.6 below.
– Moreover, if $\phi$ is lower and upper semi-continuous respectively, and $(\kappa_1,\kappa_2)$-Lipschitz in the large, then $\mathbf{T}_{t_1}^{t_2}\phi$ and $\breve{\mathbf{T}}_{t_1}^{t_2}\phi$ satisfies the following *Markov property*:

$$\mathbf{T}_{t_2}^{t_3}\circ\mathbf{T}_{t_1}^{t_2}=\mathbf{T}_{t_1}^{t_3},\quad \breve{\mathbf{T}}_{t_2}^{t_3}\circ\breve{\mathbf{T}}_{t_1}^{t_2}=\breve{\mathbf{T}}_{t_1}^{t_3},$$

whenever $t_1<t_2<t_3$. We can also have that $\lim_{t\to0^+}\mathbf{T}_0^t\phi=\phi$ and $\lim_{t\to0^+}\breve{\mathbf{T}}_0^t\phi=\phi$ if $\phi$ is lower and upper semi-continuous respectively, and $(\varepsilon,K_\varepsilon)$-Lipschitz in the large for any $\varepsilon>0$. Therefore, it is natural to set both $\mathbf{T}_t^t$ and $\breve{\mathbf{T}}_t^t$ ($t\geqslant0$) to be the identity.
– It is useful to regard the definition of $\mathbf{T}_{t_1}^{t_2}$ or $\breve{\mathbf{T}}_{t_1}^{t_2}$ as a representation of *marginal functions*. More precisely, set $F_\phi(t_1,t_2,\cdot,x,\phi(\cdot))=\phi(\cdot)+h_L(t_1,t_2,y,x,\phi(\cdot))$. If the infimum in the definition of $(\mathbf{T}_{t_1}^{t_2})\phi$ can be achieved in a compact subset $S\subset\mathbb{R}^n$, i.e.,

$$(\mathbf{T}_{t_1}^{t_2})\phi(x)=\inf_{y\in S}F_\phi(t_1,t_2,y,x,\phi(y))$$

then the Lipschitz and semiconcavity estimates can be obtained directly from the *uniform* Lipschitz and semiconcavity estimates for $h_L$ (see, for instance, [19, Theorem 3.4.4]). This is also a key point of our program for the study of the propagation of singularities of viscosity solutions (see, for instance, [13, 16, 14, 12]).

**Lemma 3.6.** *Let $t_2>t_1$ and $x\in\mathbb{R}^n$. If the function $\phi:\mathbb{R}^n\to\mathbb{R}$ is lower semi-continuous and $(\kappa_1,\kappa_2)$-Lipschitz in the large, then there exists $y\in\mathbb{R}^n$ such that $(\mathbf{T}_{t_1}^{t_2}\phi)(x)=\phi(y)+h_L(t_1,t_2,y,x,\phi(y))$. Moreover, for such a minimizer $y$ we have*

$$|y-x|\leqslant\kappa_1+\{c_0+\overline{\theta}_0(0)+\theta^*(\kappa_2+e^{2K(t_2-t_1)})+|\phi(x)|C\}(t_2-t_1)\tag{3.5}$$

*where $C=\sup_{t>0}(1-e^{-2Kt})/t$.*

*Proof.* Fix $t_2>t_1$ and $x\in\mathbb{R}^n$. For any $y\in\mathbb{R}^n$, let $\xi_y$ be a minimizer for $h_L(t_1,t_2,y,x,\phi(y))$ and $u_{\xi_y}$ is determined by

$$\begin{cases}\dot u_{\xi_y}(s)=L(s,\xi_y(s),\dot\xi_y(s),u_{\xi_y}(s)), & s\in[t_1,t_2],\\ u_{\xi_y}(t_1)=\phi(y).\end{cases}$$

It follows that

$$\dot u_{\xi_y}(s)=L(s,\xi_y(s),\dot\xi_y(s),0)+\widehat{L_u}(s)\cdot u_{\xi_y}(s),\tag{3.6}$$

where $\widehat{L_u}(s)=\int_0^1L_u(s,\xi_y(s),\dot\xi_y(s),\lambda u_{\xi_y}(s))\,d\lambda$. Solving (3.6), we obtain that

(3.7)

$$\begin{aligned}u_{\xi_y}(t_2)=&e^{\int_{t_1}^{t_2}\widehat{L_u}\,ds}\phi(y)+e^{\int_{t_1}^{t_2}\widehat{L_u}\,ds}\int_{t_1}^{t_2}e^{-\int_{t_1}^{s}\widehat{L_u}\,d\tau}L(s,\xi_y(s),\dot\xi_y(s),0)\,ds\\ \geqslant&e^{\int_{t_1}^{t_2}\widehat{L_u}\,ds}\phi(y)+\int_{t_1}^{t_2}e^{\int_s^{t_2}\widehat{L_u}\,d\tau}(\theta_0(|\dot\xi_y(s)|)-c_0)\,ds\\ \geqslant&e^{\int_{t_1}^{t_2}\widehat{L_u}\,ds}\phi(y)+e^{-K(t_2-t_1)}\int_{t_1}^{t_2}\theta_0(|\dot\xi_y(s)|)\,ds-c_0(t_2-t_1)e^{K(t_2-t_1)}.\end{aligned}$$

Let $\eta(s) \equiv x$ for $s \in [0,t]$ and $u_\eta$ satisfies

$$\begin{cases} \dot{u}_\eta(s) = L(s,\eta(s),\dot\eta(s),u_\eta(s)) = L(s,x,0,u_\eta(s)), & s \in [t_1,t_2], \\ u_\eta(t_1) = \phi(x). \end{cases}$$

Similarly, we have that

$$\begin{aligned} u_\eta(t_2) &= e^{\int_{t_1}^{t_2} \widetilde{L_u}\, ds}\phi(x) + e^{\int_{t_1}^{t_2} \widetilde{L_u}\, ds} \int_{t_1}^{t_2} e^{-\int_{t_1}^{s} \widetilde{L_u}\, d\tau} L(s,x,0,0)\, ds \\ &\leqslant e^{\int_{t_1}^{t_2} \widetilde{L_u}\, ds}\phi(x) + \int_{t_1}^{t_2} e^{\int_{s}^{t_2} \widetilde{L_u}\, d\tau}\overline{\theta}_0(0)\, ds \\ &\leqslant e^{\int_{t_1}^{t_2} \widetilde{L_u}\, ds}\phi(x) + \overline{\theta}_0(0)(t_2-t_1)e^{K(t_2-t_1)}, \end{aligned} \tag{3.8}$$

where $\widetilde{L_u}(s) = \int_0^1 L_u(s,x,0,\lambda u_\eta(s))\, d\lambda$. Combining (3.7) and (3.8) we obtain that

$$\begin{aligned} &(\phi(y) + h_L(t_1,t_2,y,x,\phi(y))) - (\phi(x) + h_L(t_1,t_2,x,x,\phi(x))) = u_{\xi_y}(t_2) - u_\eta(t_2) \\ \geqslant\, & e^{\int_{t_1}^{t_2} \widehat{L_u}\, ds}\phi(y) - e^{\int_{t_1}^{t_2} \widetilde{L_u}\, ds}\phi(x) + e^{-K(t_2-t_1)} \int_{t_1}^{t_2} \theta_0(|\dot\xi_y(s)|)\, ds \\ & - (c_0 + \overline{\theta}_0(0))(t_2-t_1)e^{K(t_2-t_1)} \\ \geqslant & -e^{K(t_2-t_1)}|\phi(y)-\phi(x)| - (e^{K(t_2-t_1)} - e^{-K(t_2-t_1)})|\phi(x)| \\ & + e^{-K(t_2-t_1)} \int_{t_1}^{t_2} \theta_0(|\dot\xi_y(s)|)\, ds - (c_0 + \overline{\theta}_0(0))(t_2-t_1)e^{K(t_2-t_1)}. \end{aligned}$$

Set $\Lambda_x = \{y \in \mathbb{R}^n : (\phi(y) + h_L(t_1,t_2,y,x,\phi(y))) - (\phi(x) + h_L(t_1,t_2,x,x,\phi(x))) \leqslant 0\}$. Notice $\Lambda_x$ is closed since $\phi$ is lower semi-continuous. Recalling that $\phi$ is $(\kappa_1,\kappa_2)$-Lipschitz in the large, for any $y \in \Lambda_x$ and any $a > 0$ we have that

$$\begin{aligned} 0 \geqslant & -e^{-K(t_2-t_1)}(\kappa_1 + \kappa_2|y-x|) - (e^{K(t_2-t_1)} - e^{-K(t_2-t_1)})|\phi(x)| \\ & + e^{-K(t_2-t_1)} \inf_{\xi \in \Gamma_{y,x}^{t_1,t_2}} \int_{t_1}^{t_2} \theta_0(|\dot\xi(s)|)\, ds - (c_0 + \overline{\theta}_0(0))(t_2-t_1)e^{K(t_2-t_1)} \\ \geqslant & -e^{-K(t_2-t_1)}(\kappa_1 + \kappa_2|y-x|) - (e^{K(t_2-t_1)} - e^{-K(t_2-t_1)})|\phi(x)| \\ & + e^{-K(t_2-t_1)} \inf_{\xi \in \Gamma_{y,x}^{t_1,t_2}} \int_{t_1}^{t_2} (a|\dot\xi(s)| - \theta^*(a))\, ds - (c_0 + \overline{\theta}_0(0))(t_2-t_1)e^{K(t_2-t_1)} \\ \geqslant & -e^{-K(t_2-t_1)}(\kappa_1 + \kappa_2|y-x|) - (e^{K(t_2-t_1)} - e^{-K(t_2-t_1)})|\phi(x)| + ae^{-K(t_2-t_1)}|y-x| \\ & - (c_0 + \overline{\theta}_0(0))(t_2-t_1)e^{K(t_2-t_1)} - \theta^*(a)(t_2-t_1)e^{-K(t_2-t_1)}. \end{aligned}$$

It follows

$$\begin{aligned} & e^{-2K(t_2-t_1)}(a-\kappa_2)|y-x| \\ \leqslant\, & \kappa_1 e^{-2K(t_2-t_1)} + |\phi(x)|C(t_2-t_1) + (c_0 + \overline{\theta}_0(0) + \theta^*(a))(t_2-t_1). \end{aligned}$$

Taking $a = \kappa_2 + e^{2K(t_2-t_1)}$, then (3.5) follows. Thus the set $\Lambda_x$ is compact and the proof is complete. □

3.2. **Representation formula.** In this section, we want to give a representation formula for the viscosity solution of (HJ) in the form of Lax-Oleinik evolution $u(t,x)$ defined as follows: for any $\phi:\mathbb{R}^n\to[-\infty,+\infty]$, set

$$\begin{aligned} u(t,x)=(\mathbf{T}_0^t\phi)(x)&=\inf_{y\in\mathbb{R}^n}\{\phi(y)+h_L(0,t,y,x,\phi(y))\}\\ &=\inf_{\xi}\left\{\int_0^t L(s,\xi,\dot{\xi},u_\xi)\,ds+\phi(\xi(0))\right\}, \end{aligned}\tag{3.9}$$

where the infimum is taken over the set

$$\mathcal{A}_{t,x}=\{\xi\in W^{1,1}([0,t],\mathbb{R}^n):\xi(t)=x\},$$

and $u_\xi$ satisfies the Carathéodory equation

$$\begin{cases}\dot{u}_\xi(s)=L(s,\xi(s),\dot{\xi}(s),u_\xi(s)), & a.e.\ s\in[0,t],\\ u_\xi(0)=\phi(\xi(0)).\end{cases}\tag{3.10}$$

The following principle of dynamic programming is analogous to the classical one.

**Proposition 3.7** (dynamic programming)**.** *Let $(t,x)\in(0,+\infty)\times\mathbb{R}^n$ and $\xi\in\mathcal{A}_{t,x}$. Then for any $t'\in[0,t]$ we have that*

$$u(t,x)\leqslant\int_{t'}^t L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds+u(t',\xi(t')),\tag{3.11}$$

*where $u_\xi$ is determined by*

$$\begin{cases}\dot{u}_\xi(s)=L(s,\xi(s),\dot{\xi}(s),u_\xi(s)), & a.e.\ s\in[t',t],\\ u_\xi(t')=u(t',\xi(t')).\end{cases}\tag{3.12}$$

*In addition, $\xi\in\mathcal{A}_{t,x}$ is a minimizer for* (3.9) *if and only if the equality holds in* (3.11) *for all $t'\in[0.t]$.*

*Proof.* Fix $t>0$ and $x\in\mathbb{R}^n$. Let $t'\in[0,t]$ and $\eta:[0,t]\to\mathbb{R}^n$ be any absolutely continuous function on $[0.t']$ such that $\eta(t')=\xi(t')$. Set

$$\gamma(s)=\begin{cases}\eta(s), & s\in[0,t'];\\ \xi(s), & s\in[t',t],\end{cases}$$

and

$$\begin{cases}\dot{u}_\gamma(s)=L(s,\gamma(s),\dot{\gamma}(s),u_\gamma(s)), & a.e.\ s\in[0,t],\\ u_\gamma(0)=\phi(\gamma(0))=\phi(\eta(0)).\end{cases}$$

It follows that

$$\begin{aligned} u_\gamma(t')&=\int_0^{t'} L(s,\gamma,\dot{\gamma},u_\gamma)\,ds+\phi(\eta(0))\\ u_\gamma(t)&=\int_{t'}^t L(s,\gamma,\dot{\gamma},u_\gamma)\,ds+u_\gamma(t')\\ &=\int_{t'}^t L(s,\gamma,\dot{\gamma},u_\gamma)\,ds+\int_0^{t'} L(s,\gamma,\dot{\gamma},u_\gamma)\,ds+\phi(\eta(0)) \end{aligned}$$

Therefore

$$u(t,x)\leqslant u_\gamma(t)\leqslant\int_{t'}^t L(s,\xi,\dot{\xi},u_\xi)\,ds+\int_0^{t'} L(s,\eta,\dot{\eta},u_\eta)\,ds+\phi(\eta(0)),$$

where $u_\xi$ and $u_\eta$ are the restriction of $u_\gamma$ on $[t,t']$ and $[0.t']$ respectively. Taking the infimum over all $\eta$ and recalling that $\xi(t')=\eta(t')$ we obtain (3.11)

Now we turn to the proof of the last assertion. If the equality holds in (3.11) for all $t'\in[0,t]$, then choosing $t'=0$ yielding that $\xi$ is a minimizer for (3.9). Conversely, if $\xi$ is a minimizer for (3.9), by (3.11) we obtain that for all $t'\in[0,t]$

$$\begin{aligned}\int_0^t L(s,\xi,\dot{\xi},u_\xi)\,ds+\phi(\xi(0))=&\,u(t,x)\\ &\leqslant\int_{t'}^t L(s,\xi,\dot{\xi},u_\xi)\,ds+u(t',\xi(t')),\end{aligned}\tag{3.13}$$

where $u_\xi$ is determined by (3.12). Invoking the definition of $u(t',\xi(t'))$, this implies the inequality in (3.13) is indeed an equality. It follows that the restriction of $\xi$ on $[0,t']$ is a minimizer for $u(t',\xi(t'))$. □

**Proposition 3.8.** *Let $\phi:\mathbb{R}^n\to\mathbb{R}$ be lower semi-continuous and $(\kappa_1,\kappa_2)$-Lipschitz in the large, and let $t>0$. Then the following holds true.*

(1) *$u(t,x)$ is finite-valued for all $t>0$ and $x\in\mathbb{R}^n$. Moreover the infimum in the definition of $u(t,x)$ is achieved by some $y_{t,x}\in\mathbb{R}^n$.*
(2) *Suppose that for any $\varepsilon>0$ there exists $K_\varepsilon>0$ such that $\phi$ is $(\varepsilon,K_\varepsilon)$-Lipschitz in the large[3]. Then $\lim_{t\to0^+}|y_{t,x}-x|=0$.*
(3) *If $\phi$ is bounded and Lipschitz with constant $\mathrm{Lip}\,(\phi)$, then there exists $\mu(t)>0$ such that $|y_{t,x}-x|\leqslant\mu(t)t$ for all $t>0$. Moreover, one can take $\mu(t)=c_0+\overline{\theta}_0(0)+\theta^*(\mathrm{Lip}\,(\phi)+e^{2Kt})+C\|\phi\|_\infty$ for some constant $C>0$.*

*Proof.* Assertion (1) is a reformulation of Lemma 3.6. For the proof of (2), set $r_\varepsilon=\overline{c}_0+\overline{\theta}_0(0)+\overline{c}_1+\theta^*(K_\varepsilon+e^{2Kt})+|\phi(x)|C$. By Lemma 3.6 we conclude

$$|y_{t,x}-x|\leqslant\varepsilon+r_\varepsilon t.$$

This implies $\lim_{t\to0^+}|y_{t,x}-x|=0$. The last assertion (3) is obvious since $\phi$ is $(0,\mathrm{Lip}\,(\phi))$-Lipschitz in the large. □

**Proposition 3.9.** *If $\phi$ is lower semi-continuous and $(\kappa_1,\kappa_2)$-Lipschitz in the large, then $u(t,x)$ defined in* (3.9) *is a viscosity solution of* (HJ).

*Remark* 3.10. Uniqueness results for (HJ) hold under further regularity assumptions. See, for instance, [5, Theorem 5.2]).

*Proof.* Fix $t>0$ and $x\in\mathbb{R}^n$. Suppose that $\varphi$ is a $C^1$-function on $(0,\infty)\times\mathbb{R}^n$ such that $u-\varphi$ attains a local maximum at $(t,x)\in U$, a neighborhood of $(t,x)$ in $(0,\infty)\times\mathbb{R}^n$. For any $(t',x')\in U$ $(t'<t)$ and any $C^1$ curve $\xi\in\Gamma^{t',t}_{x',x}$, we conclude that

$$u(t',\xi(t'))-\varphi(t',\xi(t'))\leqslant u(t,\xi(t))-\varphi(t,\xi(t)).$$

Invoking dynamic programming principle (Proposition 3.7) we obtain that

$$\frac{\varphi(t,\xi(t))-\varphi(t',\xi(t'))}{t-t'}\leqslant\frac{u(t,x)-u(t',x')}{t-t'}\leqslant\frac{1}{t-t'}\int_{t'}^t L(s,\xi,\dot{\xi},u_\xi)\,ds$$

where $u_\xi$ is determined by

$$\begin{cases}\dot{u}_\xi(s)=L(s,\xi(s),\dot{\xi}(s),u_\xi(s)), & a.e.\ s\in[t',t],\\ u_\xi(t')=u(t',\xi(t')).\end{cases}$$

[3]As mentioned in Example 3.4, a uniformly continuous function on $\mathbb{R}^n$ is $(\varepsilon,K_\varepsilon)$ Lipschitz in the large.

Taking the limit as $t' \to t$,

$$D_t\varphi(t,x) + D_x\varphi(t,x)\cdot\dot{\xi}(t) - L(t,x,\dot{\xi}(t),u(t,x)) \leqslant 0.$$

Since $\xi$ is arbitrary, we conclude

$$D_t\varphi(t,x) + H(t,x,D_x\varphi(t,x),u(t,x)) \leqslant 0.$$

This implies $u$ is viscosity subsolution of (HJ).

On the other hand, since $\phi$ is lower semi-continuous and $(\kappa_1,\kappa_2)$-Lipschitz in the large, by Proposition 3.6, there exists $y\in\mathbb{R}^n$ such that $u(t,x)=\phi(y)+h_L(0,t,y,x,\phi(y))$. Equivalently, there exists a $C^2$ curve $\xi:[0,t]\to\mathbb{R}^n$, $\xi(t)=x$, such that

$$u(t,x)=\phi(\xi(0))+\int_0^t L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds.$$

By the dynamic programming principle, we conclude that

$$u(t,x)=\int_{t'}^t L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds+u(t',\xi(t')),\quad \forall t'\in[0,t].$$

In a similar way, one can show that $u$ is viscosity supersolution of (HJ). This completes the proof. □

## Appendix A. Some facts from analysis and differential equations

A.1. **Carathéodory equations.** Let $\Omega\subset\mathbb{R}^{n+1}$ be an open set. A function $f:\Omega\subset\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ is said to satisfy *Carathéodory condition* if

- for any $x\in\mathbb{R}^n$, $f(\cdot,x)$ is measurable;
- for any $t\in\mathbb{R}$, $f(t,\cdot)$ is continuous;
- for each compact subset $U$ of $\Omega$, there is an integrable function $m_U(t)$ such that

$$|f(t,x)|\leqslant m_U(t),\quad (t,x)\in U.$$

A classical problem is to find an absolutely continuous function $x$ defined on a real interval $I$ such that $(t,x(t))\in\Omega$ for $t\in I$ and satisfies the following Carathéodory equation

$$\dot{x}(t)=f(t,x(t)),\quad a.e., t\in I. \tag{A.1}$$

**Proposition A.1** (Carathéodory)**.** *If $\Omega$ is an open set in $\mathbb{R}^{n+1}$ and $f$ satisfies the Carathéodory conditions on $\Omega$, then, for any $(t_0,x_0)$ in $\Omega$, there is a solution of* (A.1) *through* $(t_0,x_0)$*. Moreover, if the function $f(t,x)$ is also locally Lipschitzian in $x$ with a measurable Lipschitz constant, then the solution is unique.*

For the proof of Proposition A.1 and more results related to Carathéodory equation (A.1), the readers can refer to [24, 31].

A.2. **Convexity.** The following facts on the convexity is essentially known (see [17]) when the Lagrangian is independent of $t$.

**Lemma A.2.** *Let $L$ satisfy conditions* (L1)-(L3) *and $s\in[a,b]$. We conclude that*

(a) *The function*

$$f(\varepsilon):=L_v(s,x,v/(1+\varepsilon),r)\cdot v/(1+\varepsilon)-L(s,x,v/(1+\varepsilon),r) \tag{A.2}$$

*is decreasing for $\varepsilon>-1$. In particular,*

$$f(\varepsilon)\geqslant f(+\infty)=-L(s,x,0,r)\geqslant-\overline{\theta}_0(0)-K|r|.$$

(b) *If* $\varepsilon_1, \varepsilon_2 > -1$ *and* $\varepsilon_1 < \varepsilon_2$, *then we have*

$$\begin{aligned} &L(s, x, r, v/(1+\varepsilon_2)) \\ \leqslant\, &(\kappa+1)^{-1} L(s, x, r, v/(1+\varepsilon_1)) + \kappa \cdot (\kappa+1)^{-1}(\bar{\theta}_0(0) + K|r|) \end{aligned}$$

*and*

$$f(\varepsilon_2) \leqslant \kappa^{-1} L(s, x, r, v/(1+\varepsilon_1)) - (\kappa^{-1}+1) L(s, x, r, v/(1+\varepsilon_2))$$

*where* $\kappa = (\varepsilon_2 - \varepsilon_1)/(1+\varepsilon_1) > 0$.

### A.3. **Uniformly continuous functions.**

**Proposition A.3.** *Let* $f$ *be uniformly continuous function on* $\mathbb{R}^n$, *then for any* $\varepsilon > 0$ *there exists* $K > 0$ *such that*

$$|f(x) - f(y)| \leqslant K|x-y| + \varepsilon, \quad \forall x, y \in \mathbb{R}^n.$$

*Proof.* Suppose that $f$ is uniformly continuous on $\mathbb{R}^n$ and fix $\varepsilon > 0$. Then there exists $\delta > 0$ such that $|f(z) - f(z')| \leqslant \varepsilon$ whenever $|z - z'| \leqslant \delta$. For any $x, y \in \mathbb{R}^n$, let $\gamma : [0,1] \to \mathbb{R}^n$ be the straight line segment connecting $x$ to $y$, and let $\tau > 0$ such that $|x - \gamma(\tau)| = \delta$. Define $z_k = \gamma(k\tau)$, $k = 0, \ldots, N$, where $N = [\frac{1}{\tau}]$, the integer part of $\frac{1}{\tau}$. Then it is clear that

$$|z_k - z_{k+1}| = \delta,\ k = 0, \ldots, N-1, \quad \text{and} \quad |z_N - y| \leqslant \delta$$

and

$$|x-y| = \sum_{k=0}^{N-1} |z_k - z_{k+1}| + |z_N - y| \geqslant (N-1)\delta.$$

Therefore,

$$|f(x) - f(y)| \leqslant \sum_{k=0}^{N-1} |f(z_{k+1}) - f(z_k)| + |f(z_N) - f(y)| \leqslant N\varepsilon \leqslant \frac{\varepsilon}{\delta}|x-y| + \varepsilon.$$

Picking $K = \frac{\varepsilon}{\delta}$, we complete the proof. □

### A.4. **A priori estimates and existence of minimizers.**

In this section, fixing real numbers $a < b$, $u \in \mathbb{R}$, $R > 0$ and two points $x, y \in \mathbb{R}^n$ such that $|x - y| \leq R$. For convenience, we collect some a priori estimate on the minimizer $\xi$ for (1.4) and related solution $u_\xi$ of (1.3). The details of the estimates can be found in Appendix B.

We suppose $\xi$ is a minimizer for (1.4) and $u_\xi$ is determined by (1.3).

**Proposition A.4.** *There exists a continuous function* $F : [0, +\infty) \times [0, +\infty) \to [0, +\infty)$ *depending on* $R$ *and* $u$ *continuously, with* $F(r_1, \cdot)$ *being nondecreasing and superlinear and* $F(\cdot, r_2)$ *being nondecreasing, such that*

$$\begin{gathered} |u_\xi(s)| \leqslant F(b-a, R/(b-a)), \quad s \in [a,b], \\ \int_a^b |L(s, \xi(s), \dot{\xi}(s), u_\xi(s))|\, ds \leqslant F(b-a, R/(b-a)), \\ \operatorname*{ess\,inf}_{s \in [a,b]} |\dot{\xi}(s)| \leqslant F(b-a, R/(b-a)), \\ \sup_{s \in [a,b]} |\xi(s) - x| \leqslant (b-a) F(b-a, R/(b-a)). \end{gathered}$$

## APPENDIX B. EXISTENCE RESULT AND A PRIORI ESTIMATES

To just obtain an existence result for problem (1.4) under subsidiary condition (1.3), we can relax our conditions on $L$. More precisely, we suppose that $L = L(t, x, v, r) : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is a measurable function satisfying:

(L1°) $L$ is measurable in $t$, continuous in $(x, r)$ and strictly convex in $v$.
(L2°) There exist two superlinear functions $\overline{\theta}_0, \theta_0 : [0, +\infty) \to [0, +\infty)$ and two $L^\infty_{\text{loc}}$-functions $c_0, c_1 : \mathbb{R} \to [0, +\infty)$, such that

$$\overline{\theta}_0(|v|) + c_1(t) \geqslant L(t, x, v, 0) \geqslant \theta_0(|v|) - c_0(t), \quad (t, x, v) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}.$$

(L3°) There exists an $L^\infty_{\text{loc}}$-function $K : \mathbb{R} \to [0, +\infty)$ such that

$$|L(t, x, v, r) - L(t, x, v, r')| \leqslant K(t)|r - r'|, \quad r, r' \in \mathbb{R}, (t, x, v) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n.$$

Notice that $L(\cdot, \xi(\cdot), \dot{\xi}(\cdot), r)$ is measurable if $\xi$ is absolutely continuous under condition (L1°).

In this section, fixing real numbers $a < b$, $u \in \mathbb{R}$, $R > 0$ and two points $x, y \in \mathbb{R}^n$ such that $|x - y| \leqslant R$, we shall give some a priori estimate for solutions of the Carathéodory equation (1.3) under condition (L1°)-(L3°). Then we show that the action functional $J(\xi)$ defined by (1.4) attains its minimum on some element in $\Gamma^{a,b}_{x,y}$. For convenience, we set

$$u_\xi(a) = u, \qquad t = b - a.$$

Recalling Remark 1.1 we can take nonnegative constants $c_0, c_1, K$ instead of functions $c_0(\cdot), c_1(\cdot), K(\cdot)$ in our assumptions. For $\varepsilon > 0$,

$$\mathcal{A}_\varepsilon = \{\xi \in \mathcal{A} : \inf_{\eta \in \mathcal{A}} J(\eta) + \varepsilon \geq u_\xi(b) - u\}. \tag{B.1}$$

We denote $L_0(s, x, v) = L(s, x, v, 0)$ which is a Lagrangian satisfies the standard conditions in [30].

Let $\xi \in \mathcal{A}_\varepsilon$ and let $u_\xi$ be determined by

$$\dot{u}_\xi(s) = L(s, \xi(s), \dot{\xi}(s), u_\xi(s)), \quad s \in [a, b], \tag{B.2}$$

with $u_\xi(a) = u$. That is

$$\begin{aligned} \dot{u}_\xi(s) &= L_0(s, \xi(s), \dot{\xi}(s)) + (L(s, \xi(s), \dot{\xi}(s), u_\xi(s)) - L(s, \xi(s), \dot{\xi}(s), 0)) \\ &= L_0(s, \xi(s), \dot{\xi}(s)) + \widehat{L^\xi_u}(s) u_\xi(s) \end{aligned}$$

where

$$\widehat{L^\xi_u}(s) = \int_0^1 L_u(s, \xi(s), \dot{\xi}(s), \lambda u_\xi(s))\, d\lambda.$$

We remake that the function $\lambda \mapsto L_u(s, \xi(s), \dot{\xi}(s), \lambda u_\xi(s))$ and $s \mapsto \widehat{L^\xi_u}(s)$ are both bounded and measurable by condition (L3°). Therefore, Solving (B.2) we obtain

$$u_\xi(s) - u = (e^{\int_a^s \widehat{L^\xi_u}\, dr} - 1)u + \int_a^s e^{\int_\tau^s \widehat{L^\xi_u}\, dr} L_0(\tau, \xi(\tau), \dot{\xi}(\tau))\, d\tau. \tag{B.3}$$

**Lemma B.1.** *Let $\xi \in \mathcal{A}_\varepsilon$ with $u_\xi$ being determined by the associated Carathéodory equation* (1.3) *and $\varepsilon > 0$. Then there exist two continuous functions $F_1, F_2 : [0, +\infty) \times$*

$[0,+\infty) \to [0,+\infty)$ *depending on* $u$*, with* $F_i(r_1,\cdot)$ *being nondecreasing and superlinear and* $F_i(\cdot,r_2)$ *being nondecreasing for any* $r_1,r_2 \geqslant 0$*,* $i=1,2$*, such that*

$$
\begin{aligned}
&|u_\xi(s)-u| \leqslant tF_1(t,R/t)+2e^{Kt}\varepsilon, \quad s\in[a,b],\\
&\int_a^b |L(\tau,\xi(\tau),\dot\xi(\tau),u_\xi(\tau))|d\tau \leqslant F_2(t,R/t)+2e^{Kt}(1+Kt)\varepsilon.
\end{aligned} \tag{B.4}
$$

*Moreover, one can take*

$$
\begin{aligned}
F_1(r_1,r_2)=&\,3c_{r_1}e^{Kr_1}|u|+2e^{2Kr_1}(\overline{\theta}_0(r_2)+c_0),\\
F_2(r_1,r_2)=&\,2c_0+(1+Kr_1)F_1(r_1,r_2).
\end{aligned}
$$

*where* $C_t=\sup_{s\in(0,t]}\frac{e^{Ks}-1}{s}<\infty$.

*Proof.* Let $\xi\in\mathcal{A}$. By (B.3) and condition (L2°) and (L3°), we obtain that for all $s\in[a,b]$

$$
\begin{aligned}
u_\xi(s)-u\geqslant&-(e^{Kt}-1)|u|+\int_a^s e^{\int_\tau^s \widehat{L_u^\xi}\,dr}(\theta_0(|\dot\xi|)-c_0)\,d\tau\\
\geqslant&-(e^{Kt}-1)|u|-c_0\int_a^s e^{\int_\tau^s \widehat{L_u^\xi}\,dr}\,d\tau\\
\geqslant&-(e^{Kt}-1)|u|-c_0te^{Kt}.
\end{aligned}
$$

This gives the lower bound of $u_\xi$.

Now we turn to the proof of (B.4). Set $\xi_0(s)=x+s(y-x)/t$ for any $s\in[a,b]$. Then $\xi_0\in\mathcal{A}$. By solving the associated Carathéodory equation again, we have that

$$
\begin{aligned}
u_{\xi_0}(b)-u=&\,(e^{\int_a^s \widehat{L_u^{\xi_0}}\,dr}-1)u+\int_a^s e^{\int_\tau^s \widehat{L_u^{\xi_0}}\,dr}L_0(\tau,\xi_0(\tau),\dot\xi_0(\tau))\,d\tau\\
\leqslant&\,(e^{Kt}-1)|u|+te^{Kt}\overline{\theta}_0(R/t).
\end{aligned}
$$

Now, suppose $\xi\in\mathcal{A}_\varepsilon$. Then $u_\xi(b)\leqslant u_{\xi_0}(b)+\varepsilon$ and this lead to

$$
u_\xi(b)-u\leqslant(e^{Kt}-1)|u|+te^{Kt}\overline{\theta}_0(R/t)+\varepsilon. \tag{B.5}
$$

Combining the lower bound of $u_\xi$ above we obtain

$$
|u_\xi(b)-u|\leqslant(e^{Kt}-1)|u|+te^{Kt}(\overline{\theta}_0(R/t)+c_0)+\varepsilon. \tag{B.6}
$$

By (B.3) at $s=b$ we obtain

$$
u_\xi(b)-u=(e^{\int_a^b \widehat{L_u^\xi}\,dr}-1)u+\int_a^b e^{\int_\tau^b \widehat{L_u^\xi}\,dr}L_0(\tau,\xi(\tau),\dot\xi(\tau))\,d\tau.
$$

In view of (B.6) we have

$$
\begin{aligned}
&\int_a^b e^{\int_\tau^b \widehat{L_u^\xi}\,dr}|L_0(\tau,\xi(\tau),\dot\xi(\tau))|\,d\tau\leqslant|u_\xi(b)-u|+(e^{\int_a^b \widehat{L_u^\xi}\,dr}-1)|u|\\
\leqslant&\,2(e^{Kt}-1)|u|+te^{Kt}(\overline{\theta}_0(R/t)+c_0)+\varepsilon.
\end{aligned} \tag{B.7}
$$

By solving (1.3) again we have that for all $s\in[a,b]$

$$
u_\xi(b)-u=(e^{\int_s^b \widehat{L_u^\xi}\,dr}-1)u_\xi(s)+\int_s^b e^{\int_\tau^b \widehat{L_u^\xi}\,dr}L_0(\tau,\xi(\tau),\dot\xi(\tau))\,d\tau.
$$

Therefore, by (B.6) and (B.7) we conclude that for all $s \in [a,b]$

$$
\begin{aligned}
|u_\xi(s)-u_\xi(b)| \leqslant & (e^{-\int_s^b \widehat{L_u^\xi}\,dr}-1)|u_\xi(b)| + e^{-\int_s^b \widehat{L_u^\xi}\,dr}\cdot\int_s^b e^{\int_\tau^b \widehat{L_u^\xi}\,dr}|L_0(\tau,\xi(\tau),\dot{\xi}(\tau))|\,d\tau \\
\leqslant & (e^{Kt}-1)(e^{Kt}|u|+te^{Kt}(\overline{\theta}_0(R/t)+c_0)+\varepsilon) \\
& + e^{Kt}\cdot(2(e^{Kt}-1)|u|+te^{Kt}(\overline{\theta}_0(R/t)+c_0)+\varepsilon) \\
= & 3(e^{Kt}-1)e^{Kt}|u|+2te^{2Kt}(\overline{\theta}_0(R/t)+c_0)+2e^{Kt}\varepsilon
\end{aligned}
$$

which completes the proof of (B.4).

In view of (B.4) and condition (L2°) and (L3°) we have that

$$
\begin{aligned}
& \int_a^b |L_0(s,\xi,\dot{\xi})|ds \leqslant \int_a^b (L_0(s,\xi,\dot{\xi})+2c_0)\,ds \\
\leqslant & 2c_0t+u_\xi(b)-u+K\int_a^b |u_\xi|\,ds \\
\leqslant & 2c_0t+F_1(t,R/t)+2e^{Kt}\varepsilon+Kt(F_1(t,R/t)+2e^{Kt}\varepsilon) \\
= & 2c_0t+(1+Kt)F_1(t,R/t)+2e^{Kt}(1+Kt)\varepsilon.
\end{aligned}
$$

This completes the proof of the second inequality in (B.4). □

**Lemma B.2.** *Let $\xi \in \mathcal{A}_\varepsilon$ and $\varepsilon > 0$. Then there exists a continuous function $F:[0,+\infty)\times[0,+\infty)\to[0,+\infty)$ depending on u, with $F(r_1,\cdot)$ being nondecreasing and superlinear and $F(\cdot,r_2)$ being nondecreasing for any $r_1,r_2 \geqslant 0$, such that*

$$\int_0^t |\dot{\xi}(s)|\,ds \leqslant tF(t.R/t)+2e^{Kt}(1+tK)\varepsilon.$$

*Moreover, the family $\{\dot{\xi}\}_{\xi\in\mathcal{A}_\varepsilon}$ is equi-integrable.*

*Proof.* By (L2°) and (L3°), we obtain

$$
\begin{aligned}
& u_\xi(b)-u \\
= & \int_a^b L(s,\xi(s),\dot{\xi}(s),u_\xi(s))\,ds \geqslant \int_a^b \{L_0(s,\xi(s),\dot{\xi}(s))-K|u_\xi(s)|\}\,ds \\
\geqslant & \int_a^b \{\theta_0(|\dot{\xi}(s)|)-c_0-K|u_\xi(s)|\}\,ds \\
\geqslant & \int_a^b \{|\dot{\xi}(s)|-K|u_\xi(s)|-(c_0+\theta_0^*(1))\}\,ds.
\end{aligned}
\tag{B.8}
$$

In view of Lemma B.1 and (B.8), we obtain that

$$
\begin{aligned}
\int_a^b |\dot{\xi}(s)|\,ds \leqslant & \int_a^b K|u_\xi(s)|\,ds+t(c_0+\theta_0^*(1))+u_\xi(b)-u \\
\leqslant & tK(tF_1(t,R/t)+2e^{Kt}\varepsilon)+t(c_0+\theta_0^*(1)) \\
& +(tF_1(t,R/t)+2e^{Kt}\varepsilon) \\
\leqslant & tF_2(t.R/t)+2e^{Kt}(1+tK)\varepsilon.
\end{aligned}
$$

Now we turn to proof of the equi-integrability of the family $\{\dot{\xi}\}_{\xi\in\mathcal{A}_\varepsilon}$. Since $\theta_0$ is a superlinear function, for any $\alpha > 0$ there exists $C_\alpha > 0$ such that $r \leqslant \theta_0(r)/\alpha$ for

$r > C_\alpha$. Thus, for any measurable subset $E \subset [a,b]$, invoking (L2), (L3) and Lemma B.1, we have that

$$\begin{aligned}
\int_{E\cap\{|\dot\xi|>C_\alpha\}} |\dot\xi| ds &\leqslant \frac{1}{\alpha}\int_{E\cap\{|\dot\xi|>C_\alpha\}} \theta_0(|\dot\xi|) ds \leqslant \frac{1}{\alpha}\int_{E\cap\{|\dot\xi|>C_\alpha\}} \{L_0(s,\xi,\dot\xi)+c_0\}\, ds \\
&\leqslant \frac{1}{\alpha}\int_{E\cap\{|\dot\xi|>C_\alpha\}} \{L(s,\xi,\dot\xi,u_\xi)+K|u_\xi(s)|+c_0(s)\}\, ds \\
&\leqslant \frac{1}{\alpha}\int_a^b \{L(s,\xi,\dot\xi,u_\xi)+K|u_\xi(s)|+c_0(s)\}\, ds \\
&\leqslant \frac{1}{\alpha}\left\{(u_\xi(b)-u)+tK(tF_1(t,R/t)+2e^{Kt}\varepsilon+|u|)+c_0t\right\} \\
&\leqslant \frac{1}{\alpha}\left\{((tF_1(t,R/t)+2e^{Kt}\varepsilon))+tK(tF_1(t,R/t)+2e^{Kt}\varepsilon+|u|)+c_0t\right\} \\
&:= \frac{1}{\alpha}F_3(\varepsilon,t,R/t)
\end{aligned}$$

Therefore, we conclude that

$$\int_E |\dot\xi| ds \leqslant \int_{E\cap\{|\dot\xi|>C_\alpha\}} |\dot\xi| ds + \int_{E\cap\{|\dot\xi|\leqslant C_\alpha\}} |\dot\xi| ds \leqslant \frac{1}{\alpha}F_3(\varepsilon,t,R/t)+|E|C_\alpha.$$

Then, the equi-integrability of the family $\{\dot\xi\}_{\xi\in\mathcal{A}_\varepsilon}$ follows since the right-hand side can be made arbitrarily small by choosing $\alpha$ large and $|E|$ small, and this proves our claim. □

**Proposition B.3.** *The functional*

$$\mathcal{A} \ni \xi \mapsto J(\xi) = \int_a^b L(s,\xi(s),\dot\xi(s),u_\xi(s))\, ds,$$

*where $u_\xi$ is determined by* (1.3)*, admits a minimizer.*

*Remark* B.4. Notice that we can rewrite the functional $J$ as

$$J(\xi) = (e^{\int_a^b \widehat{L_u^\xi}\, dr}-1)u + \int_a^b e^{\int_\tau^b \widehat{L_u^\xi}\, dr} L_0(\tau,\xi(\tau),\dot\xi(\tau))\, d\tau \tag{B.9}$$

in spirit of (B.3) and the fact $J(\xi) = u_\xi(b)-u$. We set $\mu_\xi(s) := e^{\int_s^b \widehat{L_u^\xi}\, dr}$. Therefore $J(\xi) = J_1(\xi)+J_2(\xi)$ where

$$J_1(\xi) = (\mu_\xi(a)-1)u, \quad J_2(\xi) = \int_a^b \mu_\xi(\tau)L_0(\tau,\xi(\tau),\dot\xi(\tau))\, d\tau.$$

*Proof.* Fix $x, y \in \mathbb{R}^n$, $b > a$ and $u \in \mathbb{R}$. Consider any minimizing sequence $\{\xi_k\}$ for $J$, that is, a sequence such that $J(\xi_k) \to \inf\{J(\xi) : \xi \in \mathcal{A}\}$ as $k \to \infty$. We want to show that this sequence admits a cluster point which is the required minimizer. Notice there exists an associated sequence $\{u_{\xi_k}\}$ given by (1.3) in the definition of $J(\xi_k)$. The idea of the proof is standard but a little bit different from the classical proof of Tonelli's existence theorem.

First, notice that Lemma B.2 implies that the sequence of derivatives $\{\dot\xi_k\}$ is equi-integrable. Since the sequence $\{\dot\xi_k\}$ is equi-integrable, by the Dunford-Pettis Theorem there exists a subsequence, which we still denote by $\{\dot\xi_k\}$, and a function $\eta^* \in L^1([a,b],\mathbb{R}^n)$ such that $\dot\xi_k \rightharpoonup \eta^*$ in the weak-$L^1$ topology. The equi-integrability of $\{\dot\xi_k\}$ implies that the sequence $\{\xi_k\}$ is equi-continuous and uniformly bounded. Invoking the Ascoli-Arzela theorem, we can also assume that the sequence $\{\xi_k\}$ converges uniformly to some absolutely continuous function $\xi_\infty \in \Gamma^{a,b}_{x,y}$. For any test function $\varphi \in C^1_0([a,b],\mathbb{R}^n)$,

$$\int_a^b \varphi\eta^* ds = \lim_{k\to\infty}\int_a^b \varphi\dot{\xi}_k ds = -\lim_{k\to\infty}\int_a^b \dot{\varphi}\xi_k ds = -\int_a^b \dot{\varphi}\xi_\infty ds.$$

By the du Bois-Reymond lemma (see, for instance, [19, Lemma 6.1.1]), we conclude that $\dot{\xi}_\infty = \eta^*$ almost everywhere. In View of Remark B.4 and condition (L3), we also have that the sequence $\{\mu_{\xi_k}\}$ is bounded and equi-continuous. Therefore, $\mu_{\xi_k}$ converges uniformly to $\mu_\xi$ as $k \to \infty$ by taking a subsequence if necessary.

We recall a classical result (see, for instance, [11, Theorem 3.6] or [9, Section 3.4]) on the sequentially lower semicontinuous property on the functional

$$L^1([a,b],\mathbb{R}^m) \times L^1([a,b],\mathbb{R}^n) \ni (\alpha,\beta) \mapsto \mathbf{F}(\alpha,\beta) := \int_a^b \mathbf{L}(\alpha(s),\beta(s))\, ds.$$

One has that if (i) $\mathbf{L}$ is lower semicontinuous; (ii) $\mathbf{L}(\alpha,\cdot)$ is convex on $\mathbb{R}^n$, then the functional $\mathbf{F}$ is sequentially lower semicontinuous on the space $L^1([a,b],\mathbb{R}^m)\times L^1([a,b],\mathbb{R}^n)$ endowed with the strong topology on $L^1([a,b],\mathbb{R}^m)$ and the weak topology on $L^1([a,b],\mathbb{R}^n)$.

Now, let

$$\mathbf{L}(\mu_{\xi_k}(s),\xi_k(s),\dot{\xi}_k(s)) := \mu_{\xi_k}(s)L_0(s,\xi_k(s),\dot{\xi}_k(s))$$

with $\alpha_{\xi_k}(s) = (\mu_{\xi_k}(s),\xi_k(s))$ and $\beta_{\xi_k}(s) = \dot{\xi}_k(s)$. Then $J_2$ is lower semi-continuous in the topology mentioned above. The lower semi-continuity of $J_1$ is obvious (in fact, $J_1$ is continuous). Therefore, $\xi_\infty \in \mathcal{A}$ is a minimizer of $J$ and this completes the proof of the existence result. □

**Corollary B.5.** *There exists a continuous function $F : [0,+\infty)\times[0,+\infty) \to [0,+\infty)$ depending on $u$, with $F(r_1,r_2)$ nondecreasing in both variables and superlinear with respect to $r_2$, such that every minimizer $\xi \in \mathcal{A}$ for* (1.4) *satisfies*

$$\int_a^b |\dot{\xi}(s)|\, ds \leqslant tF(t,R/t)$$

*and*

$$\operatorname*{ess\,inf}_{s\in[a,b]} |\dot{\xi}(s)| \leqslant F(t,R/t), \qquad \sup_{s\in[a,b]} |\xi(s)-x| \leqslant tF(t,R/t).$$

*Proof.* The first assertion is a direct consequence of Lemma B.2. The last two inequalities follow from the relations

$$\operatorname*{ess\,inf}_{s\in[a,b]} |\dot{\xi}(s)| \leqslant \frac{1}{t}\int_a^b |\dot{\xi}(s)|\, ds, \quad \text{and} \quad |\xi(s)-x| \leqslant \int_a^b |\dot{\xi}(s)|\, ds,$$

together with the first assertion. □

## Appendix C. Herglotz' variational principle on manifolds

In this section, we try to explain, under the assumptions (L1)-(L4), how to move the Herglotz' generalized variational principle to a closed, connected $n$-dimensional smooth manifold $M$ without boundary. We continue to use the notations $u, t, K, c_0$ defined in Appendix B.

Once and for all, we fix a auxiliary Riemannian metric $g$ on $M$ and denote $d_g$ the distance induced by $g$. First, we notice that conditions (L1)-(L4) can be adapt to $L : \mathbb{R}\times TM\times\mathbb{R} \to \mathbb{R}$, only differences are:

– (L1) is restated as $L(t,x,\cdot,r)$ is strictly convex on $T_xM$ for any fixed $(t,x,r)$;

– the norms on $\mathbb{R}^n$ is replaced by $|\cdot|_g$ defined by $g$.

Let $\{(B_i, \Phi_i)\}$ be a $C^2$ atlas for $M$. Assume that $\{B_i\}_{i=1}^N$ is a finite open cover of $M$, where $\Phi_i : B_i \to \mathbb{D}^n$ is a $C^2$-diffeomorphism for each $i \in \{1, \dots, N\}$ and $\mathbb{D}^n$ denotes the $n$-dimensional unit disc. Thus $\Phi_j^{-1} \circ \Phi_i : B_i \cap B_j \to B_i \cap B_j$ is a $C^2$-diffeomorphism for each pair $i, j$. Let $L(t, x, v, r)$ be a Lagrangian that satisfies (L1)-(L3) together with (L4) or (L4'), for fixed $i$, let $B = B_i$ and $\Phi = \Phi_i : B \to \mathbb{D}^n$ be a corresponding local coordinate, then

$$(\Phi, d\Phi) : TB \to \mathbb{D}^n \times \mathbb{R}^n$$

defines a local trivialization of $TB$ and $L_\Phi : \mathbb{R} \times \mathbb{D}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ defined as

$$L_\Phi(t, \bar{x}, \bar{v}, u) = L(t, \Phi^{-1}(\bar{x}), d\Phi^{-1}(\bar{x})\bar{v}, u), \quad (\bar{x}, \bar{v}) \in \mathbb{D}^n \times \mathbb{R}^n,\ u \in \mathbb{R}$$

is a representation of $L$ in $(B, \Phi)$. By the local representation performed above, Herglotz' generalized variational principle for $L$ restricted on some local chart $(B, \Phi)$ is equivalent to that for $L_\Phi$ on $[a, b] \times \mathbb{D}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ if $\Phi$ is a bi-Lipschitz homeomorphism and a $C^2$-diffeomorphism.

From now on, we fix $a, b \in \mathbb{R}$. Let $x, y \in M$ and $u \in \mathbb{R}$, choosing $\xi \in \Gamma_{x,y}^{a,b}(M)$, we consider the Carathéodory equation

$$\begin{cases} \dot{u}_\xi(s) = L(s, \xi(s), \dot{\xi}(s), u_\xi(s)), & a.e.\ s \in [a, b], \\ u_\xi(a) = u. \end{cases} \tag{C.1}$$

Similarly, we define the action functional

$$J(\xi) := \int_a^b L(s, \xi(s), \dot{\xi}(s), u_\xi(s))\, ds, \tag{C.2}$$

where $\xi \in \Gamma_{x,y}^{a,b}(M)$ and $u_\xi$ is defined in (C.1). Our purpose is to minimize $J(\xi)$ over

$$\mathcal{A}(M) = \{\xi \in \Gamma_{x,y}^{a,b}(M) : \text{(C.1) admits an absolutely continuous solution } u_\xi\}.$$

Notice that $\mathcal{A}(M) \neq \varnothing$ because it contains all piecewise $C^1$ curves connecting $x$ to $y$. Moreover, (L2) implies that any $\xi \in \mathcal{A}(M)$ is absolutely continuous, thus has finite length.

For a fixed $\kappa > 0$, assume that $y \in B_{\kappa t}(x)$ and that $\eta \in \mathcal{A}(M)$ is a minimizer of the action functional $\eta \mapsto J(\eta)$. It is obvious that the estimates performed on $\mathbb{R}^n$ carry over to the manifold case, then there exist constants $C_1(\kappa, a, b) > 0$, $C(u, a, b, \kappa) > 0$ such that

$$\eta(s) \in B_{C_1 t}(x) \ \text{ for } \ s \in [a, b], \quad |\dot{\eta}(s)|_g \leqslant C_1, \quad \sup_{s \in [a,b]} |u_\eta(s)| \leqslant C. \tag{C.3}$$

The second inequality holds since we only use quantitative derivatives like $L_u, L_t$ in the deduction of Erdmann condition and it can be carried over to the manifold case.

To begin the construction, we notice that there is $r > 0$ such that for all $x \in M$, the geodesic ball $B_r(x)$ is bi-Lipschitz, $C^2$ diffeomorphic to $\mathbb{D}^n$ (some rescaling of $\Phi = \exp_x^{-1}$ shall give this diffeomorphism). We set $\kappa = \frac{\text{diam}(M)}{t}$, $C_1(\kappa, a, b) + 1 := C_2(\kappa, a, b)$.

**Local case**: Assume $2C_2 t < r$ and $x, y \in B_{\frac{r}{2}}(x_0)$ for some $x_0 \in M$. By the discussion above, Herglotz' variational principle for $L$ restricted on the local chart $(B_r(x_0), \Phi)$ is equivalent to that for $L_\Phi$, satisfying all aforementioned assumptions, defined on $[a, b] \times \mathbb{D}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. Thus, by denoting

$$\mathcal{B}(M) = \{\eta \in \mathcal{A}(M) : \eta(s) \in B_r(x) \text{ for all } s \in [a, b]\},$$

we can claim that

$$\inf_{\mathcal{A}(M)} J(\xi) = \inf_{\mathcal{B}(M)} J(\xi)$$

and they admit the same minimizers: by applying (C.3) for any minimizer $\eta$,

$$\begin{aligned} d_g(\eta(s), x_0) &\leqslant d_g(\eta(s), x) + d_g(x, x_0) \\ &\leqslant \int_a^b |\dot{\eta}(s)|_g \, ds + d_g(x, x_0) \leqslant C_1 \cdot t + \frac{r}{2} < r, \end{aligned}$$

We could formulate the conclusions from Section 2 and Appendix B into the following

**Proposition C.1.** *Assume $2C_2 t < r$ and $x, y \in B_{\frac{r}{2}}(x_0)$ for some $x_0 \in M$, $\Phi : B := B_r(x_0) \to \mathbb{D}^n$ is a local chart at $x_0$, then*

(a) *The functional*

$$\mathcal{A}(M) \ni \xi \mapsto J(\xi) = \int_a^b L(s, \xi(s), \dot{\xi}(s), u_\xi(s)) \, ds,$$

*where $u_\xi$ is determined by* (C.1) *admits a minimizer on $\mathcal{A}(B)$.*

(b) *Let $\xi \in \mathcal{A}(B)$ be a minimizer of $J$, then there is a function $F = F_{u,B} : [0, +\infty) \times [0, +\infty) \to [0, +\infty)$, with $F(\cdot, r)$ being nondecreasing for any $r \geqslant 0$, such that*

$$|u_\xi(s)| \leqslant tF(a, b, \kappa) + G(t)|u| := C(u, a, b, \kappa), \quad s \in [a, b]$$

*where $G(t) > 0$ is also nondecreasing in $t$.*

(c) *Let $\xi \in \mathcal{A}(B)$ be a minimizer of $J$, then there is a function $F = F_{u,B} : [0, +\infty) \times [0, +\infty) \to [0, +\infty)$, with $F(\cdot, r)$ is nondecreasing for any $r \geqslant 0$, such that*

$$\operatorname*{ess\,sup}_{s \in [a,b]} |\dot{\xi}(s)| \leqslant F(a, b, \kappa) := C_1(a, b, \kappa).$$

(d) *If $L$ is of class $C^2$, then for any minimizer $\xi$ for* (C.2) *we have*
   1) *Both $\xi$ and $u_\xi$ are of class $C^2$ and $\xi$ satisfies Herglotz equation* (2.18) *in local charts for all $s \in [a, b]$ where $u_\xi$ is the unique solution of* (C.1)*;*
   2) *Let $p(s) = L_v(s, \xi(s), \dot{\xi}(s), u_\xi(s))$ be the dual arc, then $p$ is also of class $C^2$ and we conclude that $(\xi, p, u_\xi)$ satisfies Lie equation* (1.6) *in local charts for all $s \in [a, b]$.*

**General case**: This is just the standard "broken geodesic" argument. Let $\{(B_i, \Phi_i)\}_{i=1}^N$ be an atlas of $M$ such that $B_i = B_{\frac{r}{2}}(x_i)$ and $\{x_i\}_{1 \leqslant i \leqslant N}$ forms a $\frac{r}{2}$-net on $M$. Without loss of generality, we assume that $x \in B_1$ and $y \in B_N$. Let $\xi \in \mathcal{A}(M)$ be a minimizer which is necessarily to be $C_1$-Lipschitz as the *a priori* estimate shown. Then, there exists a partition $a = t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = b$ such that $0 \leqslant t_{j+1} - t_j \leqslant \frac{r}{2C_2}$ and $z_j = \xi(t_j)$ and $z_{j+1} = \xi(t_{j+1})$ are contained in the same $B_i$. Thus applying Proposition C.1, as a minimizer of $J$ on $\Gamma_{z_j, z_{j+1}}^{t_j, t_{j+1}}$, $\xi|_{[t_j, t_{j+1}]}$ falls in $B_r(x_i)$, which reduce the problem to the local case. For each $j$, we define

$$h_L^j(t_j, t_{j+1}, z_j, z_{j+1}, u_j) = \inf_{\xi_j} \int_{t_j}^{t_{j+1}} L(s, \xi_j(s), \dot{\xi}_j(s), u_{\xi_j}(s)) \, ds,$$

where $\xi_j$ is an absolutely continuous curve constrained in $B_r(x_i)$ connecting $z_j$ to $z_{j+1}$ and $u_{\xi_j}$ is uniquely determined by (1.3) with initial condition $u_j$. Now we consider the problem

$$g(a,b,x,y,u) := \inf \sum_{j=1}^{k} h_L^j(t_j,t_{j+1},z_j,z_{j+1},u_j), \tag{C.4}$$

where the infimum is taken over partitions $a = t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = b$ with $t_{j+1} - t_j \in [0, \frac{r}{2C_2}]$, $z_j, z_{j+1} \in M$ contained in the some $B_i$ and $u_j \in \mathbb{R}$. Due to Proposition C.1 (b), $\{u_j\}$ can be constrained in a compact subset of $\mathbb{R}$ depending only on $u, x, y$ and $t$. Therefore the infimum in (C.4) can be attained. Thanks to the local semiconcavity of the fundamental solution $h_L^j$, $h_L^j$ is differentiable at each minimizer which leads to the fact

$$h_L(a,b,x,y,u) = g(a,b,x,y,u).$$

**Proposition C.2.** *The conclusion of Proposition C.1 holds for any connected and closed $C^2$ manifold $M$ for all $a < b$.*

## REFERENCES


[1] L. Ambrosio, O. Ascenzi, and G. Buttazzo. Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands. *J. Math. Anal. Appl.*, 142(2):301–316, 1989.

[2] M. Arisawa. Ergodic problem for the Hamilton-Jacobi-Bellman equation. I. Existence of the ergodic attractor. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 14(4):415–438, 1997.

[3] M. Arisawa. Ergodic problem for the Hamilton-Jacobi-Bellman equation. II. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 15(1):1–24, 1998.

[4] V. I. Arnol'd. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein.

[5] G. Barles. An introduction to the theory of viscosity solutions for first-order Hamilton-Jacobi equations and applications. In *Hamilton-Jacobi equations: approximations, numerical analysis and applications*, volume 2074 of *Lecture Notes in Math.*, pages 49–109. Springer, Heidelberg, 2013.

[6] P. Bernard. The Lax-Oleinik semi-group: a Hamiltonian point of view. *Proc. Roy. Soc. Edinburgh Sect. A*, 142(6):1131–1177, 2012.

[7] A. Bravetti. Contact Hamiltonian dynamics: the concept and its use. *Entropy*, 19(10):Paper No. 535, 12, 2017.

[8] A. Bravetti, H. Cruz, and D. Tapias. Contact Hamiltonian mechanics. *Ann. Physics*, 376:17–39, 2017.

[9] G. Buttazzo. *Semicontinuity, relaxation and integral representation in the calculus of variations*, volume 207 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.

[10] G. Buttazzo and M. Belloni. A survey on old and recent results about the gap phenomenon in the calculus of variations. In *Recent developments in well-posed variational problems*, volume 331 of *Math. Appl.*, pages 1–27. Kluwer Acad. Publ., Dordrecht, 1995.

[11] G. Buttazzo, M. Giaquinta, and S. Hildebrandt. *One-dimensional variational problems*, volume 15 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1998. An introduction.

[12] P. Cannarsa, Q. Chen, and W. Cheng. Dynamic and asymptotic behavior of singularities of certain weak KAM solutions on the torus. *J. Differential Equations*, 267(4):2448–2470, 2019.

[13] P. Cannarsa and W. Cheng. Generalized characteristics and Lax-Oleinik operators: global theory. *Calc. Var. Partial Differential Equations*, 56(5):Art. 125, 31, 2017.

[14] P. Cannarsa, W. Cheng, and A. Fathi. On the topology of the set of singularities of a solution to the Hamilton-Jacobi equation. *C. R. Math. Acad. Sci. Paris*, 355(2):176–180, 2017.

[15] P. Cannarsa, W. Cheng, and A. Fathi. Singularities of solutions of time dependent hamilton-jacobi equations. applications to riemannian geometry. preprint, arXiv:1912.04863, 2019.

[16] P. Cannarsa, W. Cheng, M. Mazzola, and K. Wang. Global generalized characteristics for the Dirichlet problem for Hamilton-Jacobi equations at a supercritical energy level. preprint, arXiv:1803.01591, 2018.

[17] P. Cannarsa, W. Cheng, K. Wang, and J. Yan. Herglotz' generalized variational principle and contact type Hamilton-Jacobi equations. In *Trends in control theory and partial differential equations*, volume 32 of *Springer INdAM Ser.*, pages 39–67. Springer, Cham, 2019.

[18] P. Cannarsa and M. Quincampoix. Vanishing discount limit and nonexpansive optimal control and differential games. *SIAM J. Control Optim.*, 53(4):1789–1814, 2015.
[19] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, volume 58 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA, 2004.
[20] Q. Chen, W. Cheng, H. Ishii, and K. Zhao. Vanishing contact structure problem and convergence of the viscosity solutions. *Comm. Partial Differential Equations*, 44(9):801–836, 2019.
[21] F. Clarke. A Lipschitz regularity theorem. *Ergodic Theory Dynam. Systems*, 27(6):1713–1718, 2007.
[22] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264 of *Graduate Texts in Mathematics*. Springer, London, 2013.
[23] F. H. Clarke and R. B. Vinter. Regularity properties of solutions to the basic problem in the calculus of variations. *Trans. Amer. Math. Soc.*, 289(1):73–98, 1985.
[24] E. A. Coddington and N. Levinson. *Theory of ordinary differential equations*. McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.
[25] G. Dal Maso and H. Frankowska. Autonomous integral functionals with discontinuous nonconvex integrands: Lipschitz regularity of minimizers, DuBois-Reymond necessary conditions, and Hamilton-Jacobi equations. *Appl. Math. Optim.*, 48(1):39–66, 2003.
[26] A. Davini, A. Fathi, R. Iturriaga, and M. Zavidovique. Convergence of the solutions of the discounted Hamilton-Jacobi equation: convergence of the discounted solutions. *Invent. Math.*, 206(1):29–55, 2016.
[27] A. Davini, A. Siconolfi, and M. Zavidovique. Random Lax-Oleinik semigroups for Hamilton-Jacobi systems. *J. Math. Pures Appl. (9)*, 120:294–333, 2018.
[28] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
[29] A. Fathi. Viscosity solutions of the Hamilton-Jacobi equation on a non-compact manifold. preprint, 2018.
[30] A. Fathi and E. Maderna. Weak KAM theorem on non compact manifolds. *NoDEA Nonlinear Differential Equations Appl.*, 14(1-2):1–27, 2007.
[31] A. F. Filippov. *Differential equations with discontinuous righthand sides*, volume 18 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1988. Translated from the Russian.
[32] M. Giaquinta and S. Hildebrandt. *Calculus of variations. I: The Lagrangian formalism*, volume 310 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1996.
[33] M. Giaquinta and S. Hildebrandt. *Calculus of variations. II: The Hamiltonian formalism*, volume 311 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1996.
[34] D. A. Gomes. Generalized Mather problem and selection principles for viscosity solutions and Mather measures. *Adv. Calc. Var.*, 1(3):291–307, 2008.
[35] R. B. Guenther, C. M. Guenther, and J. A. Gottsch. *The Herglotz Lectures on Contact Transformations and Hamiltonian Systems*. Juliusz Schauder Center for Nonlinear Studies. Nicholas Copernicus University, 1995.
[36] G. Herglotz. *Berührungstransformationen*. Lectures at the University of Göttingen, Göttingen, 1930.
[37] G. Herglotz. *Gesammelte Schriften*. Vandenhoeck & Ruprecht, Göttingen, 1979. With introductory articles by Peter Bergmann, S. S. Chern, Ronald B. Guenther, Claus Müller, Theodor Schneider and H. Wittich, Edited and with a foreword by Hans Schwerdtfeger.
[38] W. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.
[39] H. Ishii. The vanishing discount problem for monotone systems of Hamilton-Jacobi equations. part 1: linear coupling. preprint, arXiv:1903.00244, 2019.
[40] H. Ishii and L. Jin. The vanishing discount problem for monotone systems of Hamilton-Jacobi equations. part 2: Nonlinear coupling. preprint, arXiv:1906.07979, 2019.
[41] H. Ishii, H. Mitake, and H. V. Tran. The vanishing discount problem and viscosity Mather measures. Part 1: The problem on a torus. *J. Math. Pures Appl. (9)*, 108(2):125–149, 2017.
[42] H. Ishii, H. Mitake, and H. V. Tran. The vanishing discount problem and viscosity Mather measures. Part 2: Boundary value problems. *J. Math. Pures Appl. (9)*, 108(3):261–305, 2017.
[43] L. Jin, L. Wang, and J. Yan. A representation formula of viscosity solutions to weakly coupled systems of Hamilton-Jacobi equations with applications to regularizing effect. *J. Differential Equations*, 268(5):2012–2039, 2020.
[44] F. Legoll, M. Luskin, and R. Moeckel. Non-ergodicity of the Nosé-Hoover thermostatted harmonic oscillator. *Arch. Ration. Mech. Anal.*, 184(3):449–463, 2007.
[45] Q. Liu, P. J. Torres, and C. Wang. Contact Hamiltonian dynamics: variational principles, invariants, completeness and periodic behavior. *Ann. Physics*, 395:26–44, 2018.

[46] S. Marò and A. Sorrentino. Aubry-Mather theory for conformally symplectic systems. *Comm. Math. Phys.*, 354(2):775–808, 2017.
[47] H. Mitake, A. Siconolfi, H. V. Tran, and N. Yamada. A Lagrangian approach to weakly coupled Hamilton-Jacobi systems. *SIAM J. Math. Anal.*, 48(2):821–846, 2016.
[48] J. Moser. *Stable and random motions in dynamical systems*. Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1973. With special emphasis on celestial mechanics, Hermann Weyl Lectures, the Institute for Advanced Study, Princeton, N. J, Annals of Mathematics Studies, No. 77.
[49] S. Nosé. A unified formulation of the constant temperature molecular dynamics method. *J. Chem. Phys.*, 81:511–519, 1984.
[50] H. A. Posch, W. G. Hoover, and F. J. Vesely. Canonical dynamics of the Nosé oscillator: stability, order and chaos. *Phys. Rev. A*, 33(6):4253–4265, 1986.
[51] X. Su, L. Wang, and J. Yan. Weak KAM theory for Hamilton-Jacobi equations depending on unknown functions. *Discrete Contin. Dyn. Syst.*, 36(11):6487–6522, 2016.
[52] M. A. Sychëv. On the regularity of solutions of variational problems. *Mat. Sb.*, 183(4):118–142, 1992.
[53] K. Wang, L. Wang, and J. Yan. Implicit variational principle for contact Hamiltonian systems. *Nonlinearity*, 30(2):492–515, 2017.
[54] K. Wang, L. Wang, and J. Yan. Aubry–Mather Theory for Contact Hamiltonian Systems. *Comm. Math. Phys.*, 366(3):981–1023, 2019.
[55] K. Wang, L. Wang, and J. Yan. Variational principle for contact Hamiltonian systems and its applications. *J. Math. Pures Appl. (9)*, 123:167–200, 2019.
[56] Y.-N. Wang and J. Yan. A variational principle for contact Hamiltonian systems. *J. Differential Equations*, 267(7):4047–4088, 2019.
[57] K. Zhao and W. Cheng. On the vanishing contact structure for viscosity solutions of contact type Hamilton-Jacobi equations I: Cauchy problem. *Discrete Contin. Dyn. Syst.*, 39(8):4345–4358, 2019.

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI ROMA "TOR VERGATA", VIA DELLA RICERCA SCIENTIFICA 1, 00133 ROMA, ITALY
*E-mail address*: cannarsa@mat.uniroma2.it

DEPARTMENT OF MATHEMATICS, NANJING UNIVERSITY, NANJING 210093, CHINA
*E-mail address*: chengwei@nju.edu.cn

DEPARTMENT OF APPLIED MATHEMATICS, NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY, NANJING 210094, CHINA
*E-mail address*: jl@njust.edu.cn

SCHOOL OF MATHEMATICAL SCIENCES, SHANGHAI JIAO TONG UNIVERSITY, SHANGHAI 200240, CHINA
*E-mail address*: kzwang@sjtu.edu.cn

SCHOOL OF MATHEMATICAL SCIENCES, FUDAN UNIVERSITY AND SHANGHAI KEY LABORATORY FOR CONTEMPORARY APPLIED MATHEMATICS, SHANGHAI 200433, CHINA
*E-mail address*: yanjun@fudan.edu.cn