



# An Eikonal Equation with Vanishing Lagrangian Arising in Global Optimization

Martino Bardi<sup>1</sup> · Hicham Kouhkouh<sup>2</sup>

Accepted: 28 November 2022 / Published online: 13 March 2023  
© The Author(s) 2022

## Abstract

We show a connection between global unconstrained optimization of a continuous function  $f$  and weak KAM theory for an eikonal-type equation arising also in ergodic control. A solution  $v$  of the critical Hamilton–Jacobi equation is built by a small discount approximation as well as the long time limit of an associated evolutive equation. Then  $v$  is represented as the value function of a control problem with target, whose optimal trajectories are driven by a differential inclusion describing the gradient descent of  $v$ . Such trajectories are proved to converge to the set of minima of  $f$ , using tools in control theory and occupational measures. We prove also that in some cases the set of minima is reached in finite time.

**Keywords** Global optimization · Weak KAM theory · Exit-time control problem · Ergodic Hamilton–Jacobi equation · Occupational measures · Long time behavior of solutions · Eikonal equation · Łojasiewicz inequality

**Mathematics Subject Classification** 49L12 · 49L25 · 35Q93 · 90C26

## 1 Introduction

Let  $f \in C(\mathbb{R}^n)$  be a bounded function attaining the global minimum. Global optimization is concerned with the search of the minimum points, i.e., finding the set  $\mathfrak{M} = \operatorname{argmin} f$ . For convex smooth functions this is achieved by the gradient flow, i.e.,

---

✉ Martino Bardi  
bardi@math.unipd.it

Hicham Kouhkouh  
kouhkouh@eddy.rwth-aachen.de

<sup>1</sup> Department of Mathematics “T. Levi-Civita”, University of Padova, via Trieste, 63, 35121 Padua, Italy

<sup>2</sup> Institut für Mathematik, RTG Energy, Entropy, and Dissipative Dynamics, RWTH Aachen University, Templergraben 55 (111810), 52062 Aachen, Germany

by following the trajectories of  $\dot{y}(s) = -\nabla f(y(s))$  from any initial point  $x = y(0)$ . However, if the function  $f$  is not convex the trajectory  $y(\cdot)$  may converge to a local minimum or a saddle point. Several alternative algorithms have been designed to handle non-convex optimization, such as the stochastic gradient descent, simulated annealing, or consensus-based methods. In particular the case of non-smooth  $f$  in high dimensions is important for the applications to machine learning, see, e.g., the recent paper [14] and the references therein.

In this paper we construct and study a Lipschitz function  $v : \mathbb{R}^n \rightarrow \mathbb{R}$  such that the following normalized non-smooth gradient descent differential inclusion

$$\dot{y}(s) \in \left\{ -\frac{p}{|p|}, p \in D^-v(y(s)) \right\}, \text{ for a.e. } s > 0, \tag{1.1}$$

has a solution for any initial condition  $x = y(0)$  and all solutions converge to  $\mathfrak{M}$  as  $t \rightarrow +\infty$ . Here  $D^-v$  is the sub-differential of the theory of viscosity solutions (see, e.g., [4]). The construction of such a generating function  $v$  is based on a classical problem for Hamilton–Jacobi equations: find a constant  $c$  such that the stationary equation

$$H(x, Dv) = c \text{ in } \mathbb{R}^n, \tag{1.2}$$

has a solution  $v$ . The minimal  $c$  with this property is the critical value of the Hamiltonian  $H$  and, if  $H(x, \cdot)$  is convex, it is also the value of an optimal control problem with ergodic cost having  $H$  as its Bellman Hamiltonian. If the critical solution  $v$  is interpreted in the viscosity sense, the problem fits in the weak KAM theory, and it is well-known that, for  $H = \frac{1}{2}|p|^2 - f(x)$  with  $f$  periodic,  $c = -\min f$  [19, 28]; moreover the same holds for any bounded  $f \in C^2(\mathbb{R}^n)$  by a result of Fathi and Maderna [20], and for uniformly continuous  $f$  as proved by Barles and Roquejoffre [5]. In Sect. 2 we extend such result to  $f \in C(\mathbb{R}^n)$ , bounded, and attaining its minimum. We also prove that  $\min f$  and  $v$  solving the critical equation

$$\min f + \frac{1}{2}|\nabla v(x)|^2 = f(x) \text{ in } \mathbb{R}^n,$$

can be approximated in two ways: by the solution of the stationary equation

$$\lambda u_\lambda + \frac{1}{2}|Du_\lambda|^2 = f(x), \quad x \in \mathbb{R}^n, \tag{1.3}$$

as  $\lambda \rightarrow 0+$ , the so-called small discount limit, as well as by the long-time limit of the solution of the evolution equation

$$\partial_t u + \frac{1}{2}|Du|^2 = f(x), \text{ in } \mathbb{R}^n \times (0, +\infty), \quad u(x, 0) = 0. \tag{1.4}$$

More precisely, for the evolutive Eq. (1.4) we prove

$$\lim_{t \rightarrow +\infty} (u(x, t) - t \min f) = v(x) \quad \text{locally uniformly in } \mathbb{R}^n. \tag{1.5}$$

Note that the two problems (1.3) and (1.4) do not require the a-priori knowledge of  $\min f$  and  $\operatorname{argmin} f$ . If, in addition,  $f$  is Lipschitz and semiconcave, we show that  $v$  is semiconcave and  $Du_\lambda$  and  $D_x u(\cdot, t)$  both converge (a.e.) to  $Dv$ , therefore giving an approximation of the gradient descent Eq. (1.1). Moreover, in this case (1.1) becomes the classical normalised gradient descent

$$\dot{y}(t) = -\frac{Dv(y(t))}{|Dv(y(t))|}, \quad \forall t > 0.$$

The main result of the paper is the convergence of the gradient descent trajectories (1.1) to the set  $\mathfrak{M}$  of minima of  $f$ . This is done in Sect. 3.1 after observing that  $v$  solves also the Dirichlet problem for the eikonal equation

$$\begin{cases} |\nabla v(x)| = \ell(x), & x \in \mathbb{R}^n \setminus \mathfrak{M} \\ v(x) = 0, & x \in \mathfrak{M} \end{cases}, \tag{1.6}$$

with  $\ell(x) := \sqrt{2(f(x) - \min f)}$ . (In fact, our analysis of this problem requires only that  $\ell \in C(\mathbb{R}^n)$  is bounded, non-negative, and  $\mathfrak{M} = \{x : \ell(x) = 0\}$ ). We exploit that the unique solution of (1.6) is the value function

$$v(x) = \inf_{\alpha(\cdot)} \int_0^{t_x(\alpha)} \ell(y_x^\alpha(s)) \, ds, \quad \dot{y}_x^\alpha(s) = \alpha(s), \quad \text{for } s > 0, \quad y_x^\alpha(0) = x,$$

where  $\alpha$  is measurable,  $|\alpha(s)| \leq 1$ , and  $t_x(\alpha)$  is the first time the trajectory  $y_x^\alpha$  hits  $\mathfrak{M}$ . We show that optimal trajectories exist, satisfy the gradient descent inclusion (1.1), and tend to  $\mathfrak{M}$  as  $t \rightarrow +\infty$  under a slightly strengthened positivity condition at infinity for  $\ell$ . A crucial new tool for the proof are the occupational measures associated to these trajectories.

In the final section of the paper we give sufficient conditions such that the optimal trajectories reach  $\mathfrak{M}$  in finite time. This is a nontrivial problem even when  $v$  is smooth, because it is equivalent to the finite length of gradient orbits  $\dot{z}(s) = -Dv(z(s))$ , a question with a very large literature and open problems, see, e.g., [7, 16] and the references therein. Here we prove the finite hitting time by assuming a bound from below on  $\ell$  near the target and showing an inequality of Łojasiewicz type along optimal trajectories.

In a forthcoming companion paper we also study the approximation of  $v$  and  $\mathfrak{M}$  by vanishing viscosity. We add to (1.3) a term  $-\varepsilon \Delta u_\lambda$  and let  $\lambda \rightarrow 0+$  to get the viscous critical equation

$$U^\varepsilon - \varepsilon \Delta v^\varepsilon(x) + \frac{1}{2} |\nabla v^\varepsilon(x)|^2 = f(x) \quad \text{in } \mathbb{R}^n,$$

where  $U^\varepsilon$  is a constant. We prove that  $0 \leq U^\varepsilon - \min f \leq C\varepsilon^\beta$  for some  $\beta > 0$ . Then we define the approximate stochastic gradient descent

$$dX_s = -\nabla u_\lambda(X_s) ds + \sqrt{2\varepsilon} dW_s,$$

and show that the trajectories converge to  $\mathfrak{M}$  in a suitable sense, for small  $\lambda$  and  $\varepsilon$ . These results can be found also in the second author's thesis [27].

Note that (1.4) is the classical Hamilton–Jacobi equation with the mechanical Hamiltonian  $H(x, p) = \frac{1}{2}|p|^2 - f(x)$ , where  $-f$  is the potential energy. Then our results of Sect. 2 have an interpretation in analytical mechanics. For instance, the long-time behavior (1.5) describes a thermodynamical trend to equilibrium in a non-turbulent gas or fluid: see [12, 13].

We do not attempt to review all the literature related to the topics mentioned above. For weak KAM theory on compact manifolds we refer to [17–19], and for the PDE approach to ergodic control, mostly under periodicity assumptions, the reader can consult [1, 2] and the references therein. When the state space is not bounded one must add conditions to get some compactness. In addition to [5, 20] already quoted, such problems were studied in all  $\mathbb{R}^n$  by [3, 9, 10, 24, 30, 32] assuming that  $f$  is large enough at infinity, and by [22, 23, 25] for equations involving a linear first order term that satisfies a recurrence condition, see also the references therein. Here, instead, we get compactness from the boundedness of  $f$  and the assumption that its minimum is attained. Several of the results just quoted were used for homogenisation and singular perturbation problems, e.g., [1, 3, 28, 32], so we believe that also our results will have such applications.

The Dirichlet problem (1.6) with  $\ell$  vanishing at the boundary was studied, e.g., in [29, 31, 34]. The case of a cost that does not vanish is part of time-optimal control and it is treated in [4], see also the references therein. The synthesis of an optimal feedback from the value function  $v$  leading to (1.1) uses method from [4] based on the earlier papers [6, 21].

We do not try here to design algorithms for global optimization based on the previous results. Let us mention, however, that an efficient numerical method for computing at the same time  $c$  and  $v$  in the critical/ergodic PDE (1.2) was proposed in [8].

The paper is organized as follows. In Sect. 2.1 we prove the weak KAM theorem by the small discount approximation (1.3) and in Sect. 2.2 we study the long-time asymptotics of solutions to (1.4). Section 3.1 is devoted to the optimal control problem with target  $\mathfrak{M}$  associated to (1.6) and Sect. 3.2 to deriving the gradient descent inclusion (1.1) for the optimal trajectories. In Sect. 3.3 we prove that such trajectories converge to  $\mathfrak{M}$ , and in Sect. 3.4 we show two cases where the hitting time is finite.

## 2 A Weak KAM Theorem and Approximation of the Critical Solution

We introduce the following assumptions and refer to them wherever it is needed: *Assumptions (A)*

(A1)  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is continuous and

$$\exists \underline{f}, \overline{f} \text{ s.t. } \underline{f} \leq f(x) \leq \overline{f}, \quad \forall x \in \mathbb{R}^n. \tag{2.1}$$

(A2)  $f$  attains the minimum, i.e.,

$$\mathfrak{M} := \{x \in \mathbb{R}^n : f(x) = \underline{f} := \min_{z \in \mathbb{R}^n} f(z)\} \neq \emptyset. \tag{2.2}$$

*Assumptions (B)*

(B1)  $f$  is  $C_1$ -Lipschitz continuous, i.e.  $C_1 = \|\nabla f\|_\infty$ .

(B2)  $f$  is  $C_2$ -semiconcave, i.e.,  $D_{\xi\xi}^2 f \leq C_2$  a.e. for all  $\xi \in \mathbb{R}^n$  s.t.  $|\xi| = 1$ , where  $D_{\xi\xi}^2 f$  is the second order derivative of  $f$  in the direction  $\xi$ .

A weak KAM theorem for the Hamiltonian  $H(x, p) = \frac{1}{2}|p|^2 - f(x)$  should give conditions under which there exists a constant  $U \in \mathbb{R}$ , the (Mané) critical value, such that the equation

$$U + \frac{1}{2}|\nabla v(x)|^2 = f(x), \quad \text{in } \mathbb{R}^n, \tag{2.3}$$

has a viscosity solution  $v$ . Clearly any critical value must satisfy  $U \leq \underline{f}$ . In this section we prove under the current assumptions that  $\underline{f}$  is a critical value and construct the solution  $v$  by two different approximation procedures, both having an interpretation in terms of ergodic problems in optimal control.

The fact that  $\underline{f}$  is the maximal critical value was proved in [20] for  $f \in C^2$  and with  $\mathbb{R}^n$  replaced by any complete Riemannian manifold, by methods of weak KAM theory different from ours.

**2.1 The Small Discount Limit**

We consider the stationary approximation of (2.3)

$$\lambda u_\lambda + \frac{1}{2}|Du_\lambda|^2 = f(x), \quad x \in \mathbb{R}^n, \tag{2.4}$$

where  $\lambda > 0$  will be sent to 0. The viscosity solution  $u_\lambda$  is known to be the value function of the following infinite horizon discounted optimal control problem

$$\begin{aligned} u_\lambda(x) &= \inf_{\alpha} J(x, \alpha) := \int_0^{+\infty} \left( \frac{1}{2}|\alpha_t|^2 + f(x(t)) \right) e^{-\lambda t} dt, \\ \text{s.t. } \dot{x}(s) &= \alpha_s, \quad x(0) = x \in \mathbb{R}^n, \quad s \geq 0, \end{aligned} \tag{2.5}$$

where the controls  $\alpha : [0, +\infty) \rightarrow \mathbb{R}^n$  are measurable functions. The main result of this section is the following.

**Theorem 1** Under assumptions (A), as  $\lambda \rightarrow 0$

$$\lambda u_\lambda(x) \rightarrow \underline{f} \text{ and } u_\lambda(x) - \underline{f}\lambda^{-1} \rightarrow v(x) \text{ locally uniformly in } \mathbb{R}^n,$$

where  $v(\cdot)$  is a Lipschitz continuous viscosity solution to

$$\underline{f} + \frac{1}{2}|Dv(x)|^2 = f(x), \quad x \in \mathbb{R}^n. \tag{2.6}$$

Moreover  $v \geq 0$  in  $\mathbb{R}^n$  and null on  $\mathfrak{M}$ , and it is the unique viscosity solution of (2.6) in  $\mathbb{R}^n \setminus \mathfrak{M}$  vanishing on  $\partial\mathfrak{M}$  and bounded from below.

If we assume moreover that assumptions (B) hold, then

$$Du_\lambda(x) \rightarrow Dv(x) \text{ a.e.}$$

For the proof we need some estimates uniform in  $\lambda$ . The first Lemma is known and we omit the proof (see [27] for the details).

**Lemma 1** Under the assumption (A1), for all  $x \in \mathbb{R}^n$  and  $\lambda > 0$ ,

$$\underline{f} \leq \lambda u_\lambda(x) \leq \bar{f}, \tag{2.7}$$

$$|Du_\lambda(x)| \leq \sqrt{4\|f\|_\infty} \text{ a.e.} \tag{2.8}$$

**Lemma 2** Assume (A) and (B) hold. Then  $u_\lambda$  is  $\tilde{C}_3$ -semiconcave, where  $\tilde{C}_3$  is a positive constant independent of  $\lambda > 0$ .

**Proof** We will skip the more standard parts and refer to [27] for the complete details. We use the vanishing viscosity approximation

$$\lambda u_\lambda^\varepsilon - \varepsilon \Delta u_\lambda^\varepsilon + \frac{1}{2}|Du_\lambda^\varepsilon|^2 = f(x), \quad x \in \mathbb{R}^n. \tag{2.9}$$

We fix  $\xi \in \mathbb{R}^n$  such that  $|\xi| = 1$  and denote  $\omega_\lambda(x) := D_{\xi\xi}^2 u_\lambda^\varepsilon(x)$  the second order derivative in the direction  $\xi$ . The estimates  $\omega_\lambda(x) \leq \lambda^{-1}C_2$  and

$$|Du_\lambda^\varepsilon(x)| \leq \lambda^{-1}C_1, \tag{2.10}$$

are standard and can be got, for instance, by representing  $u_\lambda^\varepsilon$  as the value function of the stochastic infinite-horizon discounted optimal control problem associated to (2.9) and exploiting the  $C_2$ -semiconcavity and  $C_1$ -Lipschitz continuity of  $f$ .

Next we differentiate twice (2.9) in the direction of  $\xi$  and obtain

$$-\varepsilon \Delta \omega_\lambda + Du_\lambda^\varepsilon \cdot D\omega_\lambda + |D_\xi Du_\lambda^\varepsilon|^2 + \lambda \omega_\lambda = D_{\xi\xi}^2 f, \quad \text{in } \mathbb{R}^n.$$

By  $\omega_\lambda^2 \leq |D_\xi Du_\lambda|^2$  and the semiconcavity assumption  $D_{\xi\xi}^2 f \leq C_2$  we get

$$-\varepsilon \Delta \omega_\lambda + Du_\lambda^\varepsilon \cdot D\omega_\lambda + \omega_\lambda^2 + \lambda \omega_\lambda \leq C_2, \quad \text{in } \mathbb{R}^n. \tag{2.11}$$

In the case  $\omega_\lambda$  attains its maximum at some  $\bar{x}$  we have

$$\omega_\lambda^2(\bar{x}) + \lambda\omega_\lambda(\bar{x}) \leq C_2.$$

By the elementary inequality  $\frac{1}{2}(z^2 - \lambda^2) \leq z^2 + \lambda z$  we get, for  $\lambda \leq 1$ ,

$$\omega_\lambda^2(\bar{x}) \leq 2C_2 + 1,$$

and then we easily reach the conclusion. For the general case we set, for  $\beta > 0$  to be chosen,

$$\Psi_\lambda(x) := \omega_\lambda(x) - \beta \log(1 + |x|^2).$$

Since  $\omega_\lambda$  is bounded from above,  $\Psi_\lambda$  attains a global maximum in  $\mathbb{R}^n$ , say at  $\bar{x}$  (which depends on  $\lambda$  and  $\beta$ ). By evaluating (2.11) in  $\bar{x}$ , after some calculations and using the bound (2.10) we arrive at

$$\omega_\lambda^2(\bar{x}) + \lambda\omega_\lambda(\bar{x}) \leq C_2 + 2\varepsilon\beta n + 2\beta\lambda^{-1}C_1.$$

Arguing as above we get, for  $\beta \leq \lambda/2 \leq 1$ ,

$$\omega_\lambda(\bar{x})^2 \leq 2(C_1 + C_2 + 2\varepsilon n) + 1. \tag{2.12}$$

Now we claim that

$$\omega_\lambda(x) \leq C_3 := \sqrt{2(C_1 + C_2 + 2\varepsilon n) + 1}, \quad \text{for all } x \in \mathbb{R}^n.$$

To prove the claim we suppose by contradiction there exists  $y \in \mathbb{R}^n$  such that  $\omega_\lambda(y) - C_3 =: \delta > 0$ . Denote  $g(x) := \log(1 + |x|^2)$  and choose  $\beta > 0$  small enough such that  $\beta g(y) \leq \frac{\delta}{2}$ . Then

$$0 < \frac{\delta}{2} \leq \delta - \beta g(y) = \omega_\lambda(y) - \beta g(y) - C_3 = \Psi_\lambda(y) - C_3,$$

and hence  $\Psi_\lambda(\bar{x}) - C_3 > 0$ . On the other hand (2.12) gives  $\omega_\lambda(\bar{x}) \leq C_3$  and

$$\Psi_\lambda(\bar{x}) - C_3 \leq -\beta g(\bar{x}) \leq 0,$$

which is the desired contradiction. This proves the claim and the  $C_3$ -semiconcavity of  $u_\lambda^\varepsilon$ , uniformly in  $\lambda$ , for every  $0 < \varepsilon \leq 1$ . Finally we let  $\varepsilon \rightarrow 0$  in (2.9) and get that the solution  $u_\lambda$  to (2.4) is semi-concave with constant  $\tilde{C}_3 := \sqrt{2(C_1 + C_2) + 1}$ .  $\square$

**Proof of Theorem 1.** First we claim that  $\lambda u_\lambda(\bar{x}) = \underline{f}$  if  $\bar{x} \in \mathfrak{M}$  (i.e.,  $f(\bar{x}) = \underline{f} = \min f$ ), for all  $\lambda > 0$ . In fact, for such  $\bar{x}$ ,

$$u_\lambda(\bar{x}) = \inf_{\alpha} \int_0^{+\infty} \left( \frac{1}{2}|\alpha_t|^2 + f(x(t)) \right) e^{-\lambda t} dt \leq \int_0^{+\infty} f(\bar{x})e^{-\lambda t} dt = \underline{f}\lambda^{-1},$$

where the inequality follows from the choice  $\alpha \equiv 0$ . The other inequality  $\geq$  is true for all  $x \in \mathbb{R}^n$  by Lemma 1, so the claim is proved.

Now we denote  $R := \sqrt{4\|f\|_\infty}$  and use the gradient bound (2.8) to get

$$|\lambda u_\lambda(x) - \underline{f}| \leq \lambda R \text{dist}(x, \mathfrak{M}) \quad \forall x \in \mathbb{R}^n.$$

Then  $\lambda u_\lambda(x) \rightarrow \underline{f}$  locally uniformly.

Define  $\varphi_\lambda(\cdot) := u_\lambda(\cdot) - \underline{f} \lambda^{-1} \geq 0$  and use (2.8) to get, for all  $x, y \in \mathbb{R}^n$ ,

$$|\varphi_\lambda(x)| \leq R \text{dist}(x, \mathfrak{M}), \quad |\varphi_\lambda(x) - \varphi_\lambda(y)| \leq R|x - y|. \tag{2.13}$$

Hence,  $\{\varphi_\lambda(\cdot)\}_{\lambda \in (0,1)}$  is a uniformly bounded and equi-continuous family on any ball of  $\mathbb{R}^n$ . So we can choose a sequence  $\lambda_k \rightarrow 0$  as  $k \rightarrow +\infty$ , such that  $\varphi_{\lambda_k}(\cdot) \rightarrow v(\cdot) \in C(\mathbb{R}^n)$  locally uniformly. Plugging  $\varphi_\lambda$  in (2.4) we get

$$\lambda \varphi_\lambda + \underline{f} + \frac{1}{2}|D\varphi_\lambda(x)|^2 = f(x), \quad x \in \mathbb{R}^n.$$

We let  $\lambda_k \rightarrow 0$  and use the stability of viscosity solutions to find that  $v$  satisfies (2.6).

Now we note that (2.6) is an eikonal equation with right hand side  $f(x) - \underline{f} > 0$  in  $\mathbb{R}^n \setminus \mathfrak{M}$ ,  $v \geq 0$  and  $v = 0$  on  $\partial\mathfrak{M}$ . This Dirichlet boundary value problem is known to have a unique viscosity solution bounded from below. Therefore the convergence of  $\varphi_\lambda$  is for  $\lambda \rightarrow 0$  and not only on subsequences.

The convergence of the gradient  $Du_\lambda(\cdot)$  to  $Dv(\cdot)$  is a direct consequence of [11, Theorem 3.3.3], recalling that  $|\varphi_\lambda(x)| \leq R|x|$  and using the uniform semiconcavity estimate in Lemma 2. □

### 2.2 Long Time Asymptotics

Here we consider the evolutive Hamilton-Jacobi equation

$$\begin{cases} \partial_t u(x, t) + \frac{1}{2}|Du(x, t)|^2 = f(x), & (x, t) \in \mathbb{R}^n \times (0, +\infty), \\ u(x, 0) = 0, & x \in \mathbb{R}^n, \end{cases} \tag{2.14}$$

where  $D = \nabla = D_x$  denotes the gradient with respect to the space variables  $x$ , and we will study the limit as  $t \rightarrow +\infty$ . The viscosity solution  $u(x, t)$  is known to be the value function of the following finite-horizon optimal control problem

$$\begin{aligned} u(x, t) &= \inf_{\alpha} J(x, t, \alpha) := \int_0^t \frac{1}{2}|\alpha_s|^2 + f(x(s)) \, ds, \\ &\text{s.t. } \dot{x}(s) = \alpha_s, \quad x(0) = x \in \mathbb{R}^n, \end{aligned} \tag{2.15}$$



where  $\alpha : [0, +\infty) \rightarrow \mathbb{R}^n$  are measurable functions. The main result of this section is the following.

**Theorem 2** *Under assumptions (A), as  $t \rightarrow +\infty$ ,*

$$\frac{u(x, t)}{t} \rightarrow \underline{f} \text{ and } u(x, t) - \underline{f}t \rightarrow v(x) \text{ locally uniformly in } \mathbb{R}^n,$$

where  $v(\cdot)$  is the viscosity solution of (2.6) found in Theorem 1. If we assume moreover that assumptions (B) hold, then

$$D_x u(x, t) \rightarrow Dv(x) \text{ a.e.}$$

To proceed with its proof we need some estimates uniform in  $t$ .

**Lemma 3** *Under the assumption (A1), for all  $(x, t) \in \mathbb{R}^n \times (0, +\infty)$ ,*

$$\underline{f} \leq \frac{u(x, t)}{t} \leq \bar{f}, \tag{2.16}$$

$$|\partial_t u(x, t)| \leq \|f\|_\infty \text{ a.e.}, \tag{2.17}$$

$$|Du(x, t)| \leq \sqrt{4\|f\|_\infty} \text{ a.e.} \tag{2.18}$$

**Proof** The arguments are standard, for the reader’s convenience we show (2.17). Fix  $h \in \mathbb{R}$  and  $x \in \mathbb{R}^n$ . Note first that  $|u(x, h)| \leq |h|\|f\|_\infty$ . Let us now denote  $\bar{v}(x, t) := u(x, t+h) + |h|\|f\|_\infty$ . Both  $u$  and  $\bar{v}$  solve the same PDE in (2.14) with initial conditions  $u(x, 0) = 0$  and  $\bar{v}(x, 0) = u(x, h) + |h|\|f\|_\infty \geq 0$ , hence by the comparison principle in [15, Theorem 2.1] we get  $u(x, t) \leq \bar{v}(x, t)$ .

Conversely,  $\underline{v}(x, t) := u(x, t+h) - |h|\|f\|_\infty$  solves the same PDE in (2.14) with initial condition  $\underline{v}(x, 0) = u(x, h) - |h|\|f\|_\infty \leq u(x, 0) = 0$ . The same comparison principle now implies that  $\underline{v}(x, t) \leq u(x, t)$ . Therefore, one gets  $|u(x, t+h) - u(x, t)| \leq |h|\|f\|_\infty$ .  $\square$

**Lemma 4** *Assume (A) and (B) hold. Then  $u$  is  $\tilde{C}_3$ -semiconcave, where  $\tilde{C}_3$  is a positive constant independent of  $t \geq 0$ .*

**Proof** As we did in the proof of Lemma 2, we consider the vanishing viscosity approximation

$$\begin{cases} \partial_t u^\varepsilon - \varepsilon \Delta u^\varepsilon + \frac{1}{2} |\nabla u^\varepsilon|^2 = f(x), & (x, t) \in \mathbb{R}^n \times (0, \infty) \\ u^\varepsilon(x, 0) = 0, & x \in \mathbb{R}^n \end{cases} \tag{2.19}$$

It is known that  $u^\varepsilon$  is the value function of the stochastic control problem

$$u^\varepsilon(x, t) = \inf_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_0^t \frac{1}{2} |\alpha_s|^2 + f(X_s) ds \mid X_0 = x \right], \quad dX_s = \alpha_s ds + \sqrt{2\varepsilon} dW_s. \tag{2.20}$$

Take  $\xi \in \mathbb{R}^n$  with  $|\xi| = 1$  and let  $\omega(x, t) := D_{\xi\xi}^2 u^\varepsilon(x, t)$  be the second order derivative in space in the direction  $\xi$ . We claim first that  $\omega(x, t) \leq t C_2$  or, equivalently, the value function  $u^\varepsilon(x, t)$  is  $t C_2$ -semiconcave in the spatial variable  $x$ . Let  $\delta > 0$  and take a  $\frac{\delta}{2}$ -optimal control for the initial point  $x$ . By using the same control for the initial points  $x + h$  and  $x - h$  we get

$$\begin{aligned}
 & u^\varepsilon(x + h, t) - 2u^\varepsilon(x, t) + u^\varepsilon(x - h, t) - \delta \\
 & \leq \mathbb{E} \left[ \int_0^t f(X_s^{x+h}) - 2f(X_s^x) + f(X_s^{x-h}) \, ds \right].
 \end{aligned}
 \tag{2.21}$$

From the controlled diffusion in (2.20) we have  $X_s^x = \frac{1}{2} (X_s^{x+h} + X_s^{x-h})$ , and  $f$   $C_2$ -semiconcave implies

$$\begin{aligned}
 & \mathbb{E} \left[ \int_0^t f(X_s^{x+h}) - 2f(X_s^x) + f(X_s^{x-h}) \, ds \right] \\
 & \leq C_2 \mathbb{E} \left[ \int_0^t \frac{1}{4} |X_s^{x+h} - X_s^{x-h}|^2 \, ds \right] \leq t C_2 |h|^2.
 \end{aligned}
 \tag{2.22}$$

Since  $\delta > 0$  is arbitrary we have proved the claim. Similar computations (see [27]) yield

$$|Du^\varepsilon(x, t)| \leq t C_1.
 \tag{2.23}$$

Next we differentiate twice (2.19) in the direction of  $\xi$  and obtain

$$\partial_t \omega - \varepsilon \Delta \omega + Du^\varepsilon \cdot D\omega + |D_\xi Du^\varepsilon|^2 = D_{\xi\xi}^2 f, \quad \text{in } \mathbb{R}^n \times (0, T].
 \tag{2.24}$$

Since  $\omega^2 \leq |D_\xi Du^\varepsilon|^2$  and by the semiconcavity assumption  $D_{\xi\xi}^2 f \leq C_2$

$$\partial_t \omega - \varepsilon \Delta \omega + Du^\varepsilon \cdot D\omega + \omega^2 \leq C_2, \quad \text{in } \mathbb{R}^n \times (0, +\infty).
 \tag{2.25}$$

Now set  $g(x) := \log(1 + |x|^2)$  and  $\Phi(x, t) := \omega(x, t) - \beta g(x)$ , in  $\mathbb{R}^n \times (0, +\infty)$  for some  $\beta > 0$  to be made precise. Since  $\omega$  is bounded from above for  $0 \leq t \leq T$ ,  $\Phi$  admits a global maximum in  $\mathbb{R}^n \times [0, T]$ . Let  $(\bar{x}, \bar{t})$  be such a maximum point. We consider first the case  $\bar{t} \in (0, T)$  and evaluate (2.25) in  $(\bar{x}, \bar{t})$  to get

$$\omega^2(\bar{x}, \bar{t}) \leq C_2 + 2\varepsilon\beta \frac{n + (n-2)|\bar{x}|^2}{(1 + |\bar{x}|^2)^2} - 2\beta Du^\varepsilon(\bar{x}, \bar{t}) \cdot \frac{\bar{x}}{1 + |\bar{x}|^2}.
 \tag{2.26}$$

Note that  $x \in \mathbb{R}^n \mapsto \frac{n+(n-2)|x|^2}{(1+|x|^2)^2}$  has a global maximum in  $x = 0$ , and  $\frac{x}{1+|x|^2}$  is bounded. Then, by (2.23) the bound in (2.26) gives

$$\omega^2(\bar{x}, \bar{t}) \leq C_2 + 2\varepsilon\beta n + 2\beta T C_1.$$

We choose  $\beta$  and  $T$  such that  $\beta \leq 1/(2T) < 1$ . Then

$$\omega(\bar{x}, \bar{t})^2 \leq C_2 + C_1 + 2n\varepsilon. \tag{2.27}$$

On the other hand, if  $\bar{t} = 0$ ,  $u_\lambda(x, 0) = 0$  for all  $x$  implies  $\omega(\bar{x}, 0) = 0$  and (2.27) still holds. And if  $\bar{t} = T$  then  $\partial_t \Phi(\bar{x}, T) \geq 0$ , i.e.,  $\partial_t \omega(\bar{x}, T) \geq 0$  and (2.27) still holds. Therefore we have

$$\omega(\bar{x}, \bar{t}) \leq C_3 := \sqrt{C_1 + C_2 + 2\varepsilon n}. \tag{2.28}$$

We are now ready to prove that  $\omega(x, t) \leq C_3$  for all  $(x, t) \in \mathbb{R}^n \times (0, +\infty)$ . As in the proof of Theorem 1 we suppose by contradiction there exists  $(y, s)$  such  $\omega(y, s) - C_3 =: \delta > 0$ . Without loss of generality, we can choose  $T > 0$  large enough such that  $s < T$ . Then we argue exactly as in the proof of Theorem 1 and reach a contradiction by choosing  $\beta$  such that  $\beta g(y) \leq \frac{\delta}{2}$ . This proves the  $C_3$ -semiconcavity of  $u$  with respect to  $x$  uniformly in  $t$ , for every  $0 < \varepsilon \leq 1$ . Finally, we let  $\varepsilon \rightarrow 0$  in (2.19) and get that the solution  $u$  to (2.14) is semi-concave in  $x$  with constant  $\tilde{C}_3 := \sqrt{C_1 + C_2}$ .  $\square$

**Proof of Theorem 2.** First we observe that  $\frac{1}{t}u(x, t) = \underline{f}$  if  $\bar{x} \in \mathfrak{M}$ .

In fact, for such  $\bar{x}$ ,

$$u(\bar{x}, t) = \inf_{\alpha} \int_0^t \frac{1}{2} |\alpha_s|^2 + f(x(s)) \, ds \leq \int_0^t \underline{f} \, dt = t \underline{f},$$

where the inequality follows from the choice  $\alpha \equiv 0$ . The other inequality  $\geq$  is true for all  $x \in \mathbb{R}^n$  by Lemma 3.

Denote  $R := \sqrt{4\|f\|_\infty}$  and use the gradient bound (2.18) to get

$$\left| \frac{1}{t}u(x, t) - \underline{f} \right| \leq \frac{1}{t}R \operatorname{dist}(x, \mathfrak{M}) \quad \forall x \in \mathbb{R}^n, t > 0.$$

Then  $u(x, t)/t \rightarrow \underline{f}$  locally uniformly as  $t \rightarrow \infty$ .

Define now  $\varphi_t(\cdot) := u(\cdot, t) - \underline{f}t$ . We observe that, in view of (2.18),  $|\varphi_t(x)| \leq R \operatorname{dist}(x, \mathfrak{M})$  and  $|\varphi_t(x) - \varphi_t(y)| \leq R|x - y|$ . Hence,  $\{\varphi_t(\cdot)\}_{t \geq 0}$  is a locally uniformly bounded and equi-continuous family. We claim that  $\varphi_t(\cdot) \rightarrow \psi(\cdot) \in C(\mathbb{R}^n)$  locally uniformly as  $t \rightarrow +\infty$  and  $\psi(\cdot)$  is a viscosity solution of

$$\underline{f} + \frac{1}{2}|D\psi(x)|^2 = f(x), \quad \text{in } \mathbb{R}^n. \tag{2.29}$$

To prove the claim define  $u_\eta(x, t) := \varphi_{t/\eta}(x) = u\left(x, \frac{t}{\eta}\right) - \frac{t}{\eta}\underline{f}$ . Then we have

$$\eta \partial_t u_\eta + \underline{f} + \frac{1}{2}|Du_\eta|^2 = f(x), \quad \text{in } \mathbb{R}^n \times (0, \infty).$$

Now consider the upper and lower relaxed semilimits

$$\theta(x, t) := \limsup_{\eta \rightarrow 0, s \rightarrow t, y \rightarrow x} u_\eta(y, s), \quad \zeta(x, t) := \liminf_{\eta \rightarrow 0, s \rightarrow t, y \rightarrow x} u_\eta(y, s),$$

and note that they are finite by the local equiboundedness of  $\varphi_t$ . It is well-known from the stability properties of viscosity solutions (see, e.g., [4]) that they are, respectively, a sub- and supersolution of (2.29) for any  $t > 0$ . Moreover, for all  $t > 0$ ,

$$\theta(x, t) = \limsup_{s \rightarrow +\infty, y \rightarrow x} \varphi_s(y) = \limsup_{s \rightarrow +\infty} \varphi_s(x),$$

where the last equality comes from the equicontinuity of  $\varphi_t$ . Similarly,

$$\zeta(x, t) = \liminf_{s \rightarrow +\infty} \varphi_s(x),$$

and so both  $\theta$  and  $\zeta$  do not depend on  $t$ . Next note that  $\varphi_s(x) = 0$  for all  $x \in \mathfrak{M}$  and it is non-negative everywhere. Then  $\theta(x) = \zeta(x) = 0$  on  $\partial\mathfrak{M}$ , and they are a sub- and a supersolution bounded from below of (2.29) in  $\mathbb{R}^n \setminus \mathfrak{M}$ , where  $f(x) - \underline{f} > 0$ . Then a standard comparison principle for the Dirichlet problem associated to eikonal equations gives  $\theta(x) = \zeta(x)$ . This proves that  $\varphi_t$  converges pointwise to  $\psi := \theta = \zeta \geq 0$ , and the convergence is locally uniform by the Ascoli-Arzelà theorem, which gives the claim. Moreover  $\psi$  coincides with the function  $v$  found in Theorem 1.

Finally, the convergence of the gradient  $D_x u(\cdot, t) = D\varphi_t$  to  $D\psi$  is a direct consequence of [11, Theorem 3.3.3], recalling that  $|\varphi_t(x)| \leq R \operatorname{dist}(x, \mathfrak{M})$  and using the uniform semiconcavity estimate in Lemma 4. □

### 3 Reaching the Minima Via Optimal Control

#### 3.1 The Optimal Control Problem with Target

In this section we consider the Dirichlet problem

$$\begin{cases} |\nabla v(x)| = \ell(x), & x \in \mathbb{R}^n \setminus \mathfrak{M}, \\ v(x) = 0, & x \in \mathfrak{M}, \end{cases} \tag{3.1}$$

motivated by the ergodic equation (2.6) of the previous section if  $\ell(x) = \sqrt{2(f(x) - \underline{f})}$ . Here, however, the standing assumptions are only that  $\mathfrak{M} \subseteq \mathbb{R}^n$  is a closed nonempty set, possibly unbounded, and

$$\ell \in C(\mathbb{R}^n) \text{ is bounded, } \ell(x) > 0 \text{ if } x \in \mathbb{R}^n \setminus \mathfrak{M}, \ell \equiv 0 \text{ on } \mathfrak{M}. \tag{F}$$

Also define  $\bar{\ell} := \sup_{x \in \mathbb{R}^n} \ell(x)$ . The Lipschitz and semiconcavity conditions of the previous section (assumptions (B)) will not be needed in most statements of the present section.

We recall that the continuous viscosity solution of (3.1) is the value function of the control problem

$$v(x) = \inf_{\alpha} \int_0^{t_x(\alpha)} \ell(y_x^\alpha(s)) \, ds, \tag{3.2}$$

where  $\alpha$  (an admissible control) is a measurable function  $[0, +\infty) \rightarrow B(0, 1)$ , the unit ball in  $\mathbb{R}^n$ ,  $t_x(\alpha) := \inf\{s \geq 0 : y_x^\alpha(s) \in \mathfrak{M}\}$ , and

$$\dot{y}_x^\alpha(s) = \alpha(s), \quad \forall s \geq 0, \quad y_x^\alpha(0) = x. \tag{3.3}$$

**Theorem 3** *Under Assumption (F) there exists an optimal control  $\alpha^*$  for the problem (3.2).*

**Proof** Notice first that (F) allows to rewrite  $v$  as

$$v(x) = \inf \int_0^{+\infty} \ell(y_x^\alpha(s)) \, ds, \quad \text{s.t.: (3.3) with } s \mapsto \alpha(s) \in B(0, 1) \text{ measurable.}$$

Fix  $x \in \mathbb{R}^n$  and consider a minimizing sequence  $(y_k, \alpha_k)_k$ , i.e., satisfying

$$\lim_{k \rightarrow +\infty} \int_0^{+\infty} \ell(y_k(t)) \, dt = v(x), \quad y_k(t) = x + \int_0^t \alpha_k(s) \, ds, \quad \forall t \geq 0. \tag{3.4}$$

Fix  $N \in \mathbb{N}$ . Using Alaoglu’s theorem, we can extract a subsequence that we denote by  $(y_{k(N)}, \alpha_{k(N)})$ , where  $k(N) \rightarrow +\infty$ , such that

$$\begin{aligned} \alpha_{k(N)} &\overset{*}{\rightharpoonup} \alpha_N^*, \quad \text{a.e. in } [0, N], \\ y_{k(N)} &\rightarrow y_N^*, \quad \text{loc. unif. on } [0, N], \\ \text{and } y_N^*(t) &= x + \int_0^t \alpha_N^*(s) \, ds, \quad \text{for all } t \in [0, N]. \end{aligned}$$

We repeat this procedure in the interval  $[0, N + 1]$  and extract from the previous subsequence another subsequence  $(y_{k(N+1)}, \alpha_{k(N+1)})$  with the same properties in  $[0, N + 1]$ . Note that

$$\begin{aligned} \alpha_{N+1}^* &= \alpha_N^*, \quad \text{a.e. in } [0, N], \\ y_{N+1}^* &= y_N^*, \quad \text{in } [0, N]. \end{aligned}$$

This suggests the definition of the candidate optimal pair  $(y^*, \alpha^*)$  as

$$(y^*, \alpha^*) := (y_N^*, \alpha_N^*) \quad \text{in } [0, N].$$

To prove its optimality consider the diagonal subsequence  $(y_{N(N)}, \alpha_{N(N)})$ . By the previous construction, for any fixed  $T > 0$  we have

$$\begin{aligned} \alpha_{N(N)} &\xrightarrow{*} \alpha^*, \quad \text{a.e. in } [0, T], \\ y_{N(N)} &\rightarrow y^*, \quad \text{loc. unif. on } [0, T], \\ \text{and } y^*(t) &= x + \int_0^t \alpha^*(s) \, ds, \quad \text{for all } t \in [0, T]. \end{aligned} \tag{3.5}$$

Now use Fatou’s lemma

$$\int_0^\infty \liminf_{N \rightarrow +\infty} \ell(y_{N(N)}(t)) \, dt \leq \liminf_{N \rightarrow \infty} \int_0^{+\infty} \ell(y_{N(N)}(t)) \, dt.$$

By (3.4) the right-hand side is  $v(x)$  because  $y_{N(N)}$  is a subsequence of  $y_k$ . Now use the continuity of  $\ell$  in the left hand side and get

$$\int_0^\infty \ell(y^*(t)) \, dt = \int_0^\infty \liminf_{N \rightarrow +\infty} \ell(y_{N(N)}(t)) \, dt \leq v(x),$$

which says that  $(y^*, \alpha^*)$  is an optimal pair solution to (3.2). □

Next we show that the fraction of time spent by an optimal trajectory away from the minimizers of  $\ell$  tends to zero as  $t \rightarrow +\infty$ .

For a given fixed  $\delta > 0$  we define the set of quasi-minimizers

$$K_\delta := \{x \in \mathbb{R}^n : \ell(x) \leq \delta\},$$

and the fraction of time  $\rho^\delta(t)$  spent by an optimal trajectory starting from  $x$  away from  $K_\delta$

$$\rho^\delta(t) = \rho^\delta(t, x, \alpha^*) := \frac{1}{t} |\{s \in [0, t] : y_x^{\alpha^*}(s) \notin K_\delta\}|,$$

where  $|I|$  denotes the Lebesgue measure of  $I \subseteq \mathbb{R}$ . In other words,  $\rho^\delta(t)$  is the image of the complement of  $K_\delta$  by the *occupational measure* of the optimal trajectory  $y_x^{\alpha^*}$ .

**Theorem 4** *Under Assumption (F), for any  $x \in \mathbb{R}^n$  and  $\delta > 0$ , an optimal trajectory  $y_x^{\alpha^*}(\cdot)$  for the problem (3.2) satisfies*

$$\rho^\delta(t, x, \alpha^*) \leq \frac{\bar{\ell}}{t \delta} \text{dist}(x, \mathfrak{M}). \tag{3.6}$$

In particular,  $\lim_{t \rightarrow +\infty} \rho^\delta(t) = 0$ .

**Proof** Since  $\ell \geq 0$ , using the characteristic function  $\mathbb{1}_Q(y) = 1$  if  $y \in Q$  and 0 otherwise,

$$\int_0^t \ell(y_x^{\alpha^*}(s)) ds \geq \int_0^t \mathbb{1}_{K_\delta^c}(y_x^{\alpha^*}(s)) \ell(y_x^{\alpha^*}(s)) ds \geq \delta \int_0^t \mathbb{1}_{K_\delta^c}(y_x^{\alpha^*}(s)) ds,$$

and hence

$$\frac{1}{t} \int_0^t \ell(y_x^{\alpha^*}(s)) ds \geq \delta \rho^\delta(t).$$

Now, since  $\ell(y_x^{\alpha^*}(s)) = 0$  for all  $s \geq t_x(\alpha^*)$  and  $\ell(\cdot) \leq \bar{\ell}$ , we have for all  $t \geq 0$

$$\begin{aligned} \int_0^t \ell(y_x^{\alpha^*}(s)) ds &\leq \int_0^{t_x(\alpha^*)} \ell(y_x^{\alpha^*}(s)) ds, \\ &= v(x) \leq \bar{\ell} \inf \{t_x(\alpha) : (3.3) \text{ holds with } |\alpha(s)| \leq 1\}. \end{aligned}$$

The second factor on the right-hand side is the minimal time function whose optimal trajectories are the straight lines from the initial position  $x$  to its orthogonal projection on the set  $\mathfrak{M}$ , with maximal speed 1. Therefore the right-hand side in the last inequality is less or equal  $\bar{\ell}|z - x|$  for any  $z \in \mathfrak{M}$ , and then

$$v(x) \leq \bar{\ell} \text{dist}(x, \mathfrak{M}).$$

Combining the inequalities we get

$$0 \leq \delta \rho^\delta(t) \leq \frac{1}{t} \int_0^t \ell(y_x^{\alpha^*}(s)) ds \leq \frac{v(x)}{t} \leq \frac{\bar{\ell}}{t} \text{dist}(x, \mathfrak{M}),$$

which concludes the proof. □

### 3.2 A Gradient Descent Inclusion for the Optimal Trajectories

So far, we showed that an optimal control exists and the corresponding optimal trajectory does not leave the set of minimizers in average as time goes to infinity, i.e. in the sense of (3.6). We now synthesize optimal feedback controls that give the gradient descent differential inclusion anticipated in the Introduction. We recall the definition of subdifferential of a continuous function

$$D^-v(z) := \left\{ p : \liminf_{x \rightarrow z} \frac{v(x) - v(z) - p \cdot (x - z)}{|x - z|} \geq 0 \right\}.$$

**Theorem 5** Assume (F). A control  $\alpha$  with corresponding trajectory  $y(\cdot) := y_x^\alpha(\cdot)$  is optimal if and only if

$$\dot{y}(s) \in \left\{ -\frac{p}{|p|}, p \in D^-v(y(s)) \right\}, \text{ for a.e. } s \in (0, t_x(\alpha)). \tag{DI}$$

**Proof** By the dynamic programming principle, the function

$$h(t) := v(y_x^\alpha(t)) + \int_0^t \ell(y_x^\alpha(s)) ds, \quad 0 \leq t \leq t_x(\alpha), \tag{3.7}$$

is non-decreasing for all  $\alpha$ , and non-increasing (hence constant) if and only if  $\alpha$  is optimal. And since  $h$  is locally Lipschitz, we get

$$\alpha \text{ is optimal if and only if } h'(t) \leq 0 \text{ a.e. } t.$$

*Proof of Necessity* Assume  $\alpha$  is optimal, and so  $h' \leq 0$ . Let  $y(\cdot) := y_x^\alpha(\cdot)$ .

Claim 1.  $p \cdot \dot{y}(t) + \ell(y(t)) \leq 0$  for all  $p \in D^-v(y(t))$  a.e.  $t$ .

Let  $\partial^-v(x; q)$  be the lower Dini derivative at  $x$  in the direction  $q$  (see Eq. (2.47) in [4, p. 125]). Then by [4, Lemma 2.50, p. 135], one has

$$\partial^-(v \circ y)(s; 1) = \partial^-v(y(s); \dot{y}(s)),$$

and for almost every  $t$ ,  $h'(t) = \partial^-v(y(t); \dot{y}(t)) + \ell(y(t))$ . Next, using [4, Lemma 2.37, p. 126], one has, for any  $z \in \mathbb{R}^n$ ,

$$D^-v(z) = \{ p : p \cdot q \leq \partial^-v(z; q), \forall q \in \mathbb{R}^n \},$$

and hence, for almost every  $t$  and for all  $p \in D^-v(y(t))$ ,

$$p \cdot \dot{y}(t) + \ell(y(t)) \leq \partial^-v(y(t); \dot{y}(t)) + \ell(y(t)) = h'(t) \leq 0.$$

Claim 2.  $\dot{y}(t) = -\frac{p}{|p|}$  for all  $p \in D^-v(y(t))$ , a.e.  $t$ .

By [4, Proposition 5.3, p. 344],  $v$  is a bilateral supersolution of  $|Dv(x)| - \ell(x) = 0$  in  $\mathbb{R}^n \setminus \mathfrak{M}$ , i.e.  $|p| - \ell(x) = 0$  for all  $p \in D^-v(x)$ . This implies in particular that  $p \neq 0$  if  $x \notin \mathfrak{M}$ . Hence, and using claim 1 together with  $\dot{y} \in B(0, 1)$ , one gets

$$|p| = \ell(y(t)) \leq -p \cdot \dot{y}(t) \leq |p|,$$

that is,  $\dot{y}(t) = -p/|p|$ .

*Proof of Sufficiency* By the non-smooth calculus rule just recalled, for a.e.  $t$ ,

$$\begin{aligned} h'(t) &= -\partial^-v(y(t); -\dot{y}(t)) + \ell(y(t)) \\ &\leq -p \cdot (-\dot{y}(t)) + \ell(y(t)), \quad \forall p \in D^-v(y(t)). \end{aligned}$$



Then, if we assume  $y(\cdot)$  solves (DI),

$$h'(t) \leq -p \cdot \frac{p}{|p|} + \ell(y(t)) = -|p| + \ell(y(t)) \leq 0$$

because  $v$  is a supersolution of  $|Dv| - \ell = 0$  and  $p \in D^-v(y(t))$ . □

**Remark 1** Combining Theorems 3 and 5, the differential inclusion (DI) has at least a solution and all such solutions are optimal.

We recall the definition of limiting gradient of a Lipschitz function

$$D^*v(z) := \{ p : p = \lim_{n \rightarrow +\infty} Dv(x_n) \text{ for some } x_n \rightarrow z \},$$

and the super-differential of a continuous function

$$D^+v(z) := \left\{ p : \limsup_{x \rightarrow z} \frac{v(x) - v(z) - p \cdot (x - z)}{|x - z|} \leq 0 \right\}.$$

**Theorem 6** Assume (F). The following necessary and sufficient conditions of optimality hold.

(I) If  $y(\cdot)$  is optimal, then

- (i)  $\dot{y}(t) = -\frac{p}{|p|}$ , for all  $p \in D^+v(y(t))$ ,  $p \neq 0$  and almost all  $t \in (0, t_x(\alpha^*))$ ,
- (ii)  $|p| = \ell(y(t))$ , for all  $p \in D^+v(y(t))$  and all  $t \in (0, t_x(\alpha^*))$ ,
- (iii)  $D^+v(y(t))$  is a singleton for all  $t \in (0, t_x(\alpha^*))$ .
- (iv) If  $\ell(x) = \sqrt{2(f(x) - \underline{f})}$  and assumptions (A) and (B) are satisfied, then  $v$  is differentiable at all points  $y(t)$  with  $t \in (0, t_x(\alpha^*))$  and

$$\dot{y}(t) = -\frac{Dv(y(t))}{|Dv(y(t))|}, \quad \forall t \in (0, t_x(\alpha^*)). \tag{3.8}$$

(II) A sufficient condition for the optimality of  $y(\cdot)$  is

$$\dot{y}(t) \in -\left\{ \frac{p}{|p|} : p \in D^*v(y(t)) \cap D^+v(y(t)), p \neq 0 \right\}, \text{ a.e. } t. \tag{3.9}$$

**Proof** To prove (I.i) we take  $h$  defined by (3.7) and let  $\partial^+v(x; q)$  be the upper Dini derivative of  $v$  in direction  $q$ , with  $|q| = 1$ .

Claim 1.  $p \cdot \dot{y}(t) + \ell(y(t)) \leq 0$ , for all  $p \in D^*v(y(t))$ , a.e.  $t$ .

Using [4, Lemma 2.37, p. 126], one has, for any  $z \in \mathbb{R}^n$

$$D^+v(z) = \left\{ p : p \cdot q \geq \partial^+v(z; q), \forall q \in \mathbb{R}^n \right\}.$$

Hence, for  $p \in D^+v(y(t))$ , one has

$$p \cdot \dot{y}(t) + \ell(y(t)) = -p \cdot (-\dot{y}(t)) + \ell(y(t)) \leq -\partial^+v(y(t); -\dot{y}(t)) + \ell(y(t)).$$

But, as in Claim 1 in the proof of Theorem 5, and since  $y$  is optimal, one gets

$$-\partial^+v(y(t); -\dot{y}(t)) + \ell(y(t)) = h'(t) \leq 0,$$

which proves the claim.

Claim 2.  $\dot{y}(t) = -\frac{p}{|p|}$  for all  $p \in D^+v(y(t))$ ,  $p \neq 0$ , a.e.  $t$ .

Recalling  $|\dot{y}| \in B(0, 1)$  and  $v$  being a subsolution of  $|Dv| - \ell = 0$ , we have for all  $p \in D^+v(y(t))$ ,  $|p| \leq \ell(y(t)) \leq -p \cdot \dot{y}(t) \leq |p|$ , and hence, either  $p = 0$  or  $\dot{y}(t) = -\frac{p}{|p|}$ .

To prove (I.ii) we use the fact that  $h$  is non-increasing if and only if  $y(\cdot)$  is optimal. Hence, for  $t > 0$  and  $\tau > 0$  small, one has

$$\begin{aligned} h(t) - h(t - \tau) \leq 0 &\Rightarrow v(y(t)) - v(y(t - \tau)) + \int_{t-\tau}^t \ell(y(s))ds \leq 0 \\ &\Rightarrow v(y(t)) - v(y(t - \tau)) \leq -\ell(y(t))\tau + o(\tau). \end{aligned}$$

Recalling the definition of  $p \in D^+v(y(t))$ , one has

$$\begin{aligned} v(y(t)) - v(y(t - \tau)) &\geq p \cdot (y(t) - y(t - \tau)) + o(\tau) \\ \Rightarrow v(y(t)) - v(y(t - \tau)) &\geq \int_{t-\tau}^t p \cdot \alpha(s)ds + o(\tau) \geq -|p|\tau + o(\tau), \end{aligned}$$

and together with the previous inequality this yields

$$|p| \geq \ell(y(t)), \quad \forall t \in (0, t_x(\alpha^*)).$$

The other inequality is a direct consequence of  $p$  being in  $D^+v(y(t))$  and  $v$  a subsolution. This concludes the proof of statement (I.ii).

The property (I.iii) follows immediately from the equality  $|p| = \ell(y(t))$  for all  $p \in D^+v(y(t))$  and the convexity of the set  $D^+v(y(t))$ .

Under the additional conditions of (I.iv),  $v$  is semiconcave thanks to Lemma 2 (or Lemma 4).

This implies that  $v$  is differentiable at all points where the superdifferential is a singleton (see, e.g., [4, Proposition II.4.7 (c), p. 66]), and then at all  $y(t)$  with  $t \in (0, t_x(\alpha^*))$ . Hence, (DI) becomes (3.8).

To prove (II) note that at all points of differentiability of  $v$ , one has  $|Dv(z)| = \ell(z)$ . Then for all  $p \in D^*v(z)$ ,  $|p| = \ell(z)$ . And one has

$$h'(t) = \partial^+v(y(t); \dot{y}(t)) + \ell(y(t)) \leq p \cdot \dot{y}(t) + \ell(y(t)), \quad \forall p \in D^+v(y(t)).$$

Then, for  $y$  solving (3.9),  $p \neq 0$

$$h'(t) \leq -p \cdot \frac{p}{|p|} + \ell(y(t)) = -|p| + \ell(y(t)) = 0$$

which concludes the proof as it has been done for Theorem 5. □

### 3.3 Convergence of Optimal Trajectories to the Argmin

In order to show stability of  $\mathfrak{M}$ , we need an assumption which prevents  $\ell(\cdot)$  from approaching 0 when  $\text{dist}(x, \mathfrak{M}) \rightarrow \infty$ , that is,

- for all  $\delta > 0$ , there exists  $\gamma = \gamma(\delta) > 0$  such that

$$\inf\{\ell(x) : \text{dist}(x, \mathfrak{M}) > \delta\} > \gamma(\delta). \tag{H}$$

If  $\mathfrak{M}$  is bounded, then it is easy to see that this condition is equivalent to

$$\liminf_{|x| \rightarrow \infty} \ell(x) > 0,$$

which is also equivalent to Assumption (A3) in [24], Assumption (L3)–(3.2) in [10], and Assumption (L3) in [9]. The last inequality, however, is impossible when  $\mathfrak{M}$  is unbounded.

**Remark 2** An example of function with a unique global minimizer that does not satisfy hypothesis (H) is  $\ell(x) = |x|e^{-x^2}$ . In this case  $\mathfrak{M} = \{0\}$  and  $\inf\{\ell(x) : |x| > \delta\} = 0$  for all  $\delta$ .

A direct consequence of Theorem 5 is the following result.

**Corollary 1** Assume the conditions (F) and (H). Let  $y_x^{\alpha^*}(\cdot)$  be an optimal trajectory and  $\delta > 0$ . If there exists  $\tau > 0$  such that  $\text{dist}(y^*(\tau), \mathfrak{M}) > \delta$ , then, for  $\gamma(\cdot)$  defined in (H),

$$\rho^{\gamma(\delta/2)}(t) \geq \frac{\delta}{t}, \quad \forall t > \tau + \frac{\delta}{2}. \tag{3.10}$$

**Proof** Set  $y^*(\cdot) := y_x^{\alpha^*}(\cdot)$ . Since it satisfies (DI), we have  $|\dot{y}^*(\cdot)| \leq 1$  and hence  $y^*(\cdot)$  is Lipschitz continuous. Therefore, given  $\delta > 0$ , if there exists  $\tau > 0$  such that  $\text{dist}(y^*(\tau), \mathfrak{M}) > \delta$ , then

$$\begin{aligned} \delta < \text{dist}(y^*(\tau), \mathfrak{M}) &\leq \text{dist}(y^*(s), \mathfrak{M}) + |y^*(s) - y^*(\tau)| \\ &\leq \text{dist}(y^*(s), \mathfrak{M}) + |s - \tau|, \end{aligned}$$

which yields

$$\text{dist}(y^*(s), \mathfrak{M}) > \frac{\delta}{2}, \quad \forall s \in ]\tau - \delta/2, \tau + \delta/2[.$$

Hence one has

$$\ell(y^*(s)) \geq \inf \left\{ \ell(x) : \text{dist}(x, \mathfrak{M}) > \frac{\delta}{2} \right\}, \quad \forall s \in ]\tau - \delta/2, \tau + \delta/2[,$$

and together with (H), one gets

$$\ell(y^*(s)) > \gamma(\delta/2), \quad \forall s \in ]\tau - \delta/2, \tau + \delta/2[. \tag{3.11}$$

Therefore

$$\left| \{s \in [0, t] : y^*(s) \notin K_{\gamma(\delta/2)}\} \right| \geq \left| ]\tau - \delta/2, \tau + \delta/2[ \right|, \quad \forall t > \tau + \frac{\delta}{2}.$$

The latter writes as

$$t \rho^{\gamma(\delta/2)}(t) \geq \delta$$

and concludes the proof. □

We are now ready to show stability properties of the set of global minimizers  $\mathfrak{M}$  with respect to the optimal trajectories  $y_x^{\alpha^*}(\cdot)$ .

**Theorem 7** Assume (F) and (H) hold. Then for  $y^*(\cdot)$  as in (DI),

- (i)  $\mathfrak{M}$  is Lyapunov stable <sup>1</sup>,
- (ii)  $\mathfrak{M}$  is globally asymptotically stable <sup>2</sup>.

**Proof** Let  $y^*(\cdot) := y_x^{\alpha^*}(\cdot)$  be an optimal trajectory, i.e., a solution of (DI). We proceed by contradiction.

*Proof of (i).* Let  $\varepsilon > 0$  be fixed and suppose for all  $\eta > 0, \exists \tau > 0$  such that  $\text{dist}(y^*(\tau), \mathfrak{M}) > \varepsilon$  and  $\text{dist}(x, \mathfrak{M}) < \eta$ . Then from Corollary 1, one has

$$\rho^{\gamma(\varepsilon/2)}(t) \geq \frac{\varepsilon}{t}, \quad \forall t > \tau + \frac{\varepsilon}{2}.$$

And from Theorem 4, one has

$$\frac{t \gamma(\varepsilon/2)}{\ell} \rho^{\gamma(\varepsilon/2)}(t) \leq \text{dist}(x, \mathfrak{M}).$$

Therefore one gets

$$\frac{\varepsilon \gamma(\varepsilon/2)}{\ell} \leq \text{dist}(x, \mathfrak{M}),$$

<sup>1</sup> This means  $\forall \varepsilon > 0, \exists \eta > 0$  such that  $\text{dist}(x, \mathfrak{M}) \leq \eta \Rightarrow \text{dist}(y_x^{\alpha^*}(t), \mathfrak{M}) \leq \varepsilon, \forall t \geq 0$ .

<sup>2</sup> This means  $\mathfrak{M}$  is Lyapunov stable and  $\lim_{t \rightarrow +\infty} \text{dist}(y_x^{\alpha^*}(t), \mathfrak{M}) = 0$  for all  $x \in \mathbb{R}^n$ .

which contradicts  $\text{dist}(x, \mathfrak{M}) < \eta$  when we choose  $\eta < \frac{\varepsilon \gamma(\varepsilon/2)}{\bar{\ell}}$ . Hence we can conclude that, for all  $\varepsilon > 0$ , there exists  $\eta > 0$  such that if  $\text{dist}(x, \mathfrak{M}) \leq \eta$  then  $\text{dist}(y^*(t), \mathfrak{M}) \leq \varepsilon$  for all  $t$ .

*Proof of (ii).* Suppose there exist a diverging sequence  $\{\tau_k\}_{k \geq 0}$  and  $\varepsilon > 0$  such that  $\text{dist}(y^*(\tau_k), \mathfrak{M}) > \varepsilon$ . Without loss of generality, one can extract a subsequence (again denoted by  $\tau_k$ ) such that  $\tau_{k+1} - \tau_k \geq \varepsilon$ . Using Corollary 1, in particular (3.11), one has for all  $k \geq 0$

$$\ell(y^*(s)) \geq \gamma(\varepsilon/2), \quad \forall s \in ]\tau_k - \varepsilon/2, \tau_k + \varepsilon/2[$$

and therefore

$$\left| \{s \in [0, t] : y^*(s) \notin K_{\gamma(\varepsilon/2)}\} \right| > \sum_{\{k \geq 0 : \tau_k \leq t - \frac{\varepsilon}{2}\}} \left| ]\tau_k - \varepsilon/2, \tau_k + \varepsilon/2[ \right| = N(t) \varepsilon,$$

where  $N(t)$  is the number of distinct elements  $\{\tau_k\}_{k \geq 0}$  that are in  $[0, t + \varepsilon/2]$ , i.e.

$$N(t) := \#\{\tau_k : \tau_k \leq t + \varepsilon/2, k \geq 0\}.$$

The previous inequality writes as

$$t \rho^{\gamma(\varepsilon/2)}(t) > N(t) \varepsilon.$$

On the other hand, we know from Theorem 4, in particular (3.6), that

$$t \rho^{\gamma(\varepsilon/2)}(t) \leq \frac{\bar{\ell} \text{dist}(x, \mathfrak{M})}{\gamma(\varepsilon/2)},$$

and so we have  $N(t) < \frac{\bar{\ell} \text{dist}(x, \mathfrak{M})}{\varepsilon \gamma(\varepsilon/2)}$ . But this cannot be true since  $N(t) \rightarrow +\infty$  as  $t \rightarrow +\infty$ , and hence it concludes the proof. □

### 3.4 On Reaching the Argmin in Finite Time

Here we investigate whether the hitting time  $t_x(\alpha^*)$  of an optimal trajectory with the target  $\mathfrak{M}$  is finite or not. In view of the gradient descent inclusion (1.1), or its smooth version (3.8), the question is equivalent to the finite length of the orbits of the gradient flow  $\dot{y} \in -D^-v(y)$ , or  $\dot{y} = -\nabla v(y)$ . This is a classical problem with a large literature. Positive results require strong regularity of  $v$ , such as quasiconvexity and subanalyticity [7]. On the other hand, counterexamples are known for  $v \in C^\infty(\mathbb{R}^2)$  and target a circle [33] or a single point [16].

In our case  $v$  is not smooth, but it is the value function of a control problem and solves an eikonal equation. These properties can be exploited to prove that the hitting time is finite in some cases.

The first sufficient condition, that complements the hypothesis (H), is the following, where  $d(x) := \text{dist}(x, \mathfrak{M})$ :

- there exist a continuous function  $\tilde{\gamma}(s) > 0$  for all  $s > 0$  and  $\tilde{\gamma}(0) = 0$ , and some  $r > 0$  such that

$$\ell(x) = \tilde{\gamma}(d(x)), \quad \forall x \text{ s.t. } d(x) \leq r. \tag{L}$$

**Proposition 1** Assume (F), (H), and (L) hold, and  $\alpha^*$  be an optimal control for problem (3.2). Then the hitting time  $t_x(\alpha^*) = d(x)$  whenever  $d(x) \leq r$  and it is finite for all  $x$ .

**Proof** Let us first note that the finiteness for all  $x$  follows from the property in the case  $d(x) \leq r$ , because by Theorem 7 (ii) there exists a finite time  $\tilde{t}_x$  such that  $d(y_x^{\alpha^*}(\tilde{t}_x)) \leq r$ .

We assume that the initial position  $x$  satisfies  $d(x) \leq r$  and aim to prove that

$$v(x) = \int_0^{d(x)} \tilde{\gamma}(s) \, ds, \tag{3.12}$$

where  $v(x)$  is the value function defined in (3.2). Denote by  $V(x)$  the right-hand side of the last equality.

We first claim that  $v(x) \leq V(x)$ . Take  $z$  is in the set of projections of  $x$  onto  $\mathfrak{M}$  and consider the straight line from  $x$  to  $z$  given by the trajectory  $\bar{y}_x(t) = x - pt$ ,  $t \geq 0$ , where  $p = \frac{x-z}{|x-z|}$ . Note that  $\bar{t}_x := \inf\{t \geq 0 : \bar{y}_x(s) \in \mathfrak{M}\} = d(x)$ , and that  $d(x - pt) \leq r$  for all  $0 \leq t \leq \bar{t}_x$ . Then, by (L),

$$v(x) \leq \int_0^{\bar{t}_x} \ell(\bar{y}_x(t)) \, dt = \int_0^{\bar{t}_x} \tilde{\gamma}(d(\bar{y}_x(t))) \, dt =: J(x).$$

Observe now that  $d(\bar{y}_x(t)) = |x - z| - t = d(x) - t$ . Therefore, using the change of variable  $s := d(\bar{y}_x(t)) = d(x) - t$ , we obtain

$$J(x) = \int_0^{d(x)} \tilde{\gamma}(d(\bar{y}_x(t))) \, dt = \int_0^{d(x)} \tilde{\gamma}(s) \, ds = V(x),$$

and this proves the claim.

Next we show that  $v(x) \geq V(x)$ . Since  $v(x)$  is a continuous viscosity solution to (3.1), then using [34, Theorem 3.2 (ii)] it satisfies the upper optimality principle [34, Definition 3.1], that is,

$$v(x) \geq \inf_{\alpha} \int_0^t \ell(y_x^{\alpha}(s)) \, ds + v(y_x^{\alpha}(t)), \quad \forall t \geq 0,$$

where the dynamics of  $y_x^{\alpha}(\cdot)$  is again (3.3) with  $|\alpha(s)| \leq 1$ . Using (L) and  $v \geq 0$  we get

$$v(x) \geq \inf_{\alpha} \int_0^t \tilde{\gamma}(d(y_x^{\alpha}(s))) \, ds, \quad \forall t \geq 0.$$

In particular, since  $\tilde{\gamma}(s) = 0$  if and only if  $s = 0$ , we have

$$v(x) \geq \inf_{\alpha \in B(0,1)} \int_0^{t_x(\alpha)} \tilde{\gamma}(d(y_x^\alpha(s))) \, ds =: W(x).$$

Then the function  $W(x)$  solves in the viscosity sense the Dirichlet problem

$$\begin{cases} |\nabla W(x)| = \tilde{\gamma}(d(x)), & x \in \mathbb{R}^n \setminus \mathfrak{M} \\ W(x) = 0, & x \in \mathfrak{M}. \end{cases} \tag{3.13}$$

But  $V(x) := \int_0^{d(x)} \tilde{\gamma}(s) \, ds$  is also a viscosity solution of this Dirichlet problem because  $|D^\pm V(x)| = |D^\pm d(x)|\tilde{\gamma}(d(x))$ . We conclude using [29, Theorem 1 and Remark 3.1] that  $V(x) = W(x)$  and hence  $v(x) \geq V(x)$ .

Finally we use in the integral of the formula (3.12) the same change of variable as above to get

$$v(x) = \int_0^{d(x)} \tilde{\gamma}(d(\bar{y}_x(t))) \, dt = \int_0^{d(x)} \ell(\bar{y}_x(t)) \, dt.$$

This proves that  $\bar{y}_x(t) := x - pt$  is an optimal trajectory and  $d(x)$  is its hitting time. □

**Remark 3** In some control problems it may happen that an optimal trajectory remains arbitrarily close to a target without ever reaching it. Such a behavior has been observed in a linear-quadratic control problem studied in [26, Sect. 6.1] with the target is a singleton  $\{x_o\}$  and the time  $t_\epsilon$  of being  $\epsilon$ -close to  $x_o$  is shown to be  $t_\epsilon = C \ln\left(\frac{|x-x_o|}{\epsilon}\right)$ , where  $x$  is the initial state. Moreover, an optimal trajectory oscillates periodically around  $x_o$  (see [26, p. 55]).

Next we show that, under the set of assumptions of Sect. 2, a bound from below on  $\ell$  near the target is a sufficient condition for the finite hitting time. The proof uses an inequality of Łojasiewicz type along optimal gradient orbits.

**Theorem 8** Assume  $\ell(x) = \sqrt{2(f(x) - \underline{f})}$ , (A), (B), and (H) are satisfied, and for some  $c, r > 0, 0 < \beta < 3/2$ ,

$$\ell(x) \geq c d(x)^\beta, \quad \forall x \text{ s.t. } d(x) \leq r. \tag{3.14}$$

If  $\alpha^*$  is an optimal control for  $x$ , then the hitting time  $t_x(\alpha^*)$  is finite for all  $x$ , and for  $d(x)$  sufficiently small

$$t_x(\alpha^*) \leq \frac{C}{1 - 2\beta/3} d(x)^{\frac{3}{2}-\beta}. \tag{3.15}$$

**Proof** Set  $y(t) := y_x^{\alpha^*}(t)$  and recall from Theorem 7 that

$$\lim_{t \rightarrow t_x(\alpha^*)} d(y(t)) = 0.$$

Therefore it is not restrictive to assume that  $d(y(t)) \leq r$  for all  $t > 0$ .

We re-parametrise the trajectory  $y$  to get a gradient orbit. Set

$$s(t) := \int_0^t |Dv(y(\tau))|^{-1} d\tau \in [0, T), \quad 0 \leq t < t_x(\alpha^*),$$

where  $T \leq +\infty$ . Define  $s \mapsto t(s)$ ,  $[0, T) \rightarrow [0, t_x(\alpha^*))$ , the inverse function of  $s(t)$  and  $z(s) := y(t(s))$ . Then

$$\dot{z}(s) = -Dv(z(s)), \quad z(0) = x, \quad \lim_{s \rightarrow T} d(z(s)) = 0,$$

and

$$t(s) = \int_0^s |Dv(z(\tau))| d\tau = \int_0^s |\dot{z}(\tau)| d\tau.$$

Therefore

$$t_x(\alpha^*) = \lim_{s \rightarrow T} t(s) = \int_0^T |\dot{z}(\tau)| d\tau,$$

and so  $t_x(\alpha^*) < \infty$  if the length of the gradient orbit  $z(\cdot)$  is finite. By Theorem 6,  $v$  is differentiable at all points  $z(s)$ ,  $s > 0$ , and then

$$|Dv(z(s))| = \ell(z(s)) \geq c d(z(s))^\beta, \quad \forall s > 0, \tag{3.16}$$

by (3.14) and  $d(z(s)) \leq r$ . On the other hand, by assumptions (A2) and (B1), for some  $C_3 > 0$

$$\ell(x) \leq C_3 \sqrt{d(x)}.$$

By repeating the 1st half of the proof of Proposition 1 we get

$$v(x) \leq \int_0^{d(x)} C_3 \sqrt{s} ds = \frac{2C_3}{3} d(x)^{3/2}. \tag{3.17}$$

By combining this with (3.16) we obtain

$$|Dv(z(s))| \geq C_4 v(z(s))^\rho,$$



where  $\rho := 2\beta/3 < 1$ . This is a Łojasiewicz inequality along the gradient orbit  $z(\cdot)$ , and we can use the following classical argument:

$$\frac{-1}{1-\rho} \frac{d}{ds} [v(z(s))^{1-\rho}] = \frac{-Dv(z(s)) \cdot \dot{z}(s)}{v(z(s))^\rho} = \frac{|Dv(z(s))| |\dot{z}(s)|}{v(z(s))^\rho} \geq C_4 |\dot{z}(s)|,$$

which integrated from 0 to  $T$  gives

$$t_x(\alpha^*) \leq \frac{v(x)^{1-\rho}}{C_4(1-\rho)}.$$

Now we combine this with (3.17) to get the estimate (3.15).  $\square$

**Acknowledgements** The authors wish to thank Piermarco Cannarsa and Olivier Ley for useful conversations and the referees for their careful reading and insightful remarks.

**Funding** Martino Bardi is member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM). Hicham Kouhkhouch is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 320021702/GRK2326—Energy, Entropy, and Dissipative Dynamics (EDDY). The results of this paper are part of his Ph.D. thesis [27] which was conducted when he was a Ph.D. student in University of Padova.

## Declarations

**Conflict of interest** The authors have not disclosed any competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Alvarez, O., Bardi, M.: Ergodicity, Stabilization, and Singular Perturbations for Bellman–Isaacs Equations. Mem. Amer. Math. Soc., Providence (2010)
2. Arisawa, M., Lions, P.-L.: On ergodic stochastic control. Commun. Part. Differ. Equ. **23**, 2187–2217 (1998)
3. Artstein, Z., Gaitsgory, V.: The value function of singularly perturbed control systems. Appl. Math. Optim. **41**, 425–445 (2000)
4. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Birkhauser, Boston (1997)
5. Barles, G., Roquejoffre, J.-M.: Ergodic type problems and large time behaviour of unbounded solutions of Hamilton–Jacobi equations. Commun. Part. Differ. Equ. **31**, 1209–1225 (2006)
6. Berkovitz, L.D.: Optimal feedback controls. SIAM J. Control Optim. **27**, 991–1006 (1989)
7. Bolte, J., Daniilidis, A., Lewis, A.: The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Control Optim. **17**, 1205–1223 (2007)
8. Cacace, S., Camilli, F.: A generalized Newton method for homogenization of Hamilton–Jacobi equations. SIAM J. Sci. Comput. **38**, A3589–A3617 (2016)

9. Cannarsa, P., Mendico, C.: Asymptotic analysis for Hamilton–Jacobi equations associated with sub-Riemannian control systems. Preprint at <http://arxiv.org/abs/2012.09099> (2020)
10. Cannarsa, P., Mendico, C.: Asymptotic analysis for Hamilton–Jacobi–Bellman equations on Euclidean space. *J. Differ. Equ.* **332**, 83–122 (2022)
11. Cannarsa, P., Sinestrari, C.: *Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control*, vol. 58. Birkhauser, Boston (2004)
12. Cardin, F.: Fluid Dynamical Features of the Weak KAM Theory, in *Waves and Stability in Continuous Media*, pp. 108–117. World Scientific, Singapore (2008)
13. Cardin, F.: *Elementary Symplectic Topology and Mechanics*. Springer, Berlin (2015)
14. Carrillo, J.A., Jin, S., Li, L., Zhu, Y.: A consensus-based global optimization method for high dimensional machine learning problems, *ESAIM: Control. Optim. Calculus Var.* **27**, S5 (2021)
15. Da Lio, F., Ley, O.: Uniqueness results for second-order Bellman–Isaacs equations under quadratic growth assumptions and applications. *SIAM J. Control Optim.* **45**, 74–106 (2006)
16. Daniilidis, A., Ley, O., Sabourau, S.: Asymptotic behaviour of self-contracted planar curves and gradient orbits of convex functions. *J. Math. Pures et Appl.* **94**, 183–199 (2010)
17. Evans, L.C.: A survey of partial differential equations methods in weak KAM theory. *Commun. Pure Appl. Math.* **57**, 445–480 (2004)
18. Fathi, A.: Théoreme KAM faible et théorie de Mather sur les systemes Lagrangiens. *Comptes Rendus de l’Acad. Sci.-Ser. I-Math.* **324**, 1043–1046 (1997)
19. Fathi, A.: *Weak KAM Theorem in Lagrangian Dynamics*. Version 10. Cambridge University Press, Cambridge (2008)
20. Fathi, A., Maderna, E.: Weak KAM theorem on non compact manifolds. *Nonlinear Differ. Equ. Appl.* **14**, 1–27 (2007)
21. Frankowska, H.: Optimal trajectories associated with a solution of the contingent Hamilton–Jacobi equation. *Appl. Math. Optim.* **19**, 291–311 (1989)
22. Fujita, Y., Ishii, H., Loreti, P.: Asymptotic solutions of Hamilton–Jacobi equations in Euclidean  $n$  space. *Indiana Univ. Math. J.* **55**, 1671–1700 (2006)
23. Ishii, H.: Asymptotic solutions for large time of Hamilton–Jacobi equations in Euclidean  $n$  space. *Ann. l’IHP Anal. Non linéaire* **25**, 231–266 (2008)
24. Ishii, H., Siconolfi, A.: The vanishing discount problem for Hamilton–Jacobi equations in the Euclidean space. *Commun. Part. Differ. Equ.* **45**, 525–560 (2020)
25. Kaise, H., Sheu, S.-J.: Ergodic type Bellman equations of first order with quadratic Hamiltonian. *Appl. Math. Optim.* **59**, 37–73 (2009)
26. Kouhkhoh, H.: *Dynamic Programming Interpretation of Turnpike and Hamilton–Jacobi–Bellman Equation*. Master thesis, Paris-Saclay University (2018)
27. Kouhkhoh, H.: *Some Asymptotic Problems for Hamilton–Jacobi–Bellman Equations and Applications to Global Optimization*. PhD thesis, University of Padova (2022)
28. Lions, P.-L., Papanicolaou, G., Varadhan, S.R.: *Homogenization of Hamilton–Jacobi Equations*. Unpublished preprint (1987)
29. Malisoff, M.: Bounded-from-below solutions of the Hamilton–Jacobi equation for optimal control problems with exit times: vanishing Lagrangians, eikonal equations, and shape-from-shading. *Nonlinear Differ. Equ. Appl.* **11**, 95–122 (2004)
30. Motta, M., Sartori, C.: Asymptotic problems in optimal control with a vanishing Lagrangian and unbounded data. *Discret. Contin. Dyn. Syst.* **35**, 4527 (2015)
31. Motta, M., Sartori, C.: The value function of an asymptotic exit-time optimal control problem. *Nonlinear Differ. Equ. Appl.* **22**, 21–44 (2015)
32. Nguyen, T., Siconolfi, A.: Singularly perturbed control systems with noncompact fast variable. *J. Differ. Equ.* **261**, 4593–4630 (2016)
33. Palis, J.J., De Melo, W.: *Geometric Theory of Dynamical Systems: An Introduction*. Springer, Berlin (1982)
34. Soravia, P.: Optimality principles and representation formulas for viscosity solutions of Hamilton–Jacobi equations. I. Equations of unbounded and degenerate control problems without uniqueness. *Adv. Differ. Equ.* **4**, 275–296 (1999)