Electronic Journal of Differential Equations, Vol. 2006(2006), No. ??, pp. 1–21.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu (login: ftp)

BOUNDARY-VALUE PROBLEMS FOR HAMILTONIAN SYSTEMS AND ABSOLUTE MINIMIZERS IN CALCULUS OF VARIATIONS

VASSILI N. KOLOKOL'TSOV, ALEXEY E. TYUKOV

Abstract. We apply the method of Hamilton shooting to obtain the well-posedness of boundary value problems for certain Hamiltonian systems and some estimates for their solutions. The examples of Hamiltonian functions covered by the method include elliptic polynomials and exponentially growing functions. As a consequence we prove global existence, smoothness and almost everywhere uniqueness of absolute minimizers in the corresponding problem of calculus of variations and hence construct the global field of extremals.

1. Introduction

The classical problem of calculus of variations consists in finding a curve $\widetilde{y}(\tau)$ connecting x_0 and x in time t such that

$$I_{\widetilde{y}}(t,x,x_0) = \min_{y(0)=x_0, y(t)=x} I_y(t,x,x_0), \tag{1.1}$$

where

$$I_y(t,x,x_0) = \int_0^t L(y(\tau), \dot{y}(\tau))\, d\tau. \tag{1.2}$$

Here the function $L : \mathbb{R}^{2d} \to \mathbb{R}$ is called the Lagrangian of problem (1.1). The value of minimum (1.1) is called the *two-point function* corresponding to the Lagrangian L and will be denoted $S(t,x,x_0)$. It is usually assumed to be convex with respect to the second variable. The function

$$H(x,p) = \sup_{v\in\mathbb{R}^d} (pv - L(x,v)) \tag{1.3}$$

is called the Hamiltonian of problem (1.1). The celebrated Tonelli theorem is known to give the existence of $\widetilde{y}(\tau)$ under mild assumptions on L. The proof is based on the use of the so called direct methods of calculus of variations. However the minimizer $\widetilde{y}(\tau)$ given by Tonelli's theorem may be singular (see [2] for an example and discussion). The aim of our paper is to single out some general enough class of Lagrangians (or Hamiltonians) having always non-singular minimizers and moreover,

2000 *Mathematics Subject Classification.* 49J99, 34B15.
Key words and phrases. Hamiltonian systems; absolute minimizers; global field of extremals; WKB method.

Submitted ??. Published ??.

to prove the existence of the global field of smooth extremals for these classes. For a review on existence of non-singular minimizers see e.g. [3], [8].

In this paper we apply the method of Hamilton shooting to obtain existence, uniqueness of smooth solutions of the boundary value problems for systems

$$\begin{aligned} \dot{x} &= H_p(x,p) \\ \dot{p} &= -H_x(x,p) \end{aligned} \tag{1.4}$$

with a rather general class of Hamiltonians. This class of Hamiltonians includes the elliptic polynomials and functions of uniform exponential growth. In turn the well-posedness of boundary value problems and uniform estimates for the domain of uniqueness for boundary value problems imply the global existence, uniqueness and smoothness of absolute minimizers for problem (1.1) and hence the existence of a global field of extremals. It is proved that the smooth solutions always exist and are unique almost everywhere i.e. for any $x_0 \in \mathbb{R}^d, t > 0$ the set of those $x \in \mathbb{R}^d$ for which a trajectory delivering the absolute minimum to functional (1.2) is not unique is a closed set of Lebesgue measure zero. Our method also yields estimates for the "two-point function" $S(t, x, x_0)$.

As an important application of our results, let us mention the construction of local and global fields of extremals. It is well-known that the construction of global field of extremals corresponding to a Hamiltonian H is the crucial step in the construction of WKB-type asymptotic for solutions of pseudo-differential equations

$$i\,\frac{\partial u}{\partial t} = H\big(x, i\,\frac{\partial}{\partial x}\big)\,u \quad \text{and} \quad \frac{\partial u}{\partial t} = H\big(x, -\frac{\partial}{\partial x}\big)\,u$$

(see e.g. [4], [7]).

The method of Hamiltonian shooting can be applied also to some degenerate (non strictly convex) Hamiltonians. For example in [4] a rather general class of degenerate quadratic in momentum Hamiltonians was introduced (called regular in [4]) for which one can prove not only local uniqueness and global existence of solutions but also one can obtain exact asymptotic expansions of the two-point function that are quite similar to the case of nondegenerate quadratic Hamiltonians. It is worth mentioning that while developing the theory of stochastic Hamilton-Jacobi equations we used the method of Hamiltonian shooting to construct solutions for stochastic Hamiltonian systems driven by semimartingale noise (see [4, 5, 6]).

2. Main results

We denote by $\mathcal{F}$ the set of functions $h : \mathbb{R} \to \mathbb{R}$ represented by a series $h(z) = \sum_{n=0}^{\infty} a_n z^n$ with the radius of convergence equal to infinity such that for some constant $M = M(h) > 0$ and all $z \geq M$

$$|h^{(n)}(z)h^{(m)}(z)| \leq |h^{(n-1)}(z)h^{(m+1)}(z)| \quad \forall\, n > m \geq 0\,. \tag{2.1}$$

Here $h^{(n)}(z)$ is the n-the derivative of $h(z)$. In particular, class $\mathcal{F}$ contains functions $h_1(z) = \exp\{z\}$ and $h_2(z) = z^k$, $k \in \mathbb{N}$.

Notation. We write $(X(t), P(t)) = (X(t, x_0, p_0), P(t, x_0, p_0))$ for the solution of the system (1.4) with initial conditions (x_0, p_0) at $t = 0$ and call the x-projection $X(t)$ of a solution $(X(t), P(t))$ characteristic of the system (1.4). Denote by

$$\widetilde{x}(\tau) = \widetilde{x}(\tau; t, x, x_0) \tag{2.2}$$

a characteristic of (1.4) with boundary conditions $\widetilde{x}(0) = x_0$, $\widetilde{x}(t) = x$. We will use these notations throughout the paper.

Theorem 2.1. *Let $H \in C^4(\mathbb{R}^{2d})$, $d \geq 1$. We assume that for some $h \in \mathcal{F}$ and $\varepsilon > 0$ the following conditions hold:*

(i) $H_{pp}(x,p) \geq \varepsilon\big(1+|h''(|p|)|\big)\, E_d$ *for all $x, p \in \mathbb{R}^d$, where $E_d \in \mathbb{R}^{d\times d}$ is identity matrix;*

(ii)
$$\sum_{|I|\leq 3, |J|=k} \Big|\frac{\partial^{|I|+|J|} H(x,p)}{\partial x^I \partial p^J}\Big| \leq |h^{(k)}(|p|)|\chi_\varepsilon(p) + \varepsilon^{-1}(1-\chi_\varepsilon(p))$$
for all $x, p \in \mathbb{R}^d$ such that $|p| \geq \varepsilon^{-1}$, where
$$\chi_\varepsilon(p) = \begin{cases} 1 & \text{if } |p| \geq \varepsilon^{-1} \\ 0 & \text{otherwise} \end{cases},$$
$I = (i_1, \dots, i_{|I|})$, $J = (j_1, \dots, j_{|J|})$ *are multi-indexes,* $k = 0, 1, 2, 3$.

Then there exist $r > 0$ and $T > 0$ such that for any $x_0 \in \mathbb{R}^d$, $x \in B_r(x_0) = \{y : |y - x_0| \leq r\}$ and $t \in (0, T)$ there exists $p_0 = p_0(t, x, x_0) \in \mathbb{R}^d$ such that
$$X(t, x_0, p_0(t, x, x_0)) \equiv x\,. \tag{2.3}$$
Moreover, $p_0(t, x, x_0)$ is continuously differentiable with respect to all variables.

In other words the theorem claims that the boundary value problem for the system (1.4) is well-posed in a neighborhood of any $x_0 \in \mathbb{R}^d$, i.e. for small $|x - x_0|$ and small t there exists the unique solution of the system (1.4) such that $x(0) = x_0$, $x(t) = x$.

The lengthy assumptions of Theorem 2.1 are designed to include the main examples of Hamiltonian functions $H(x,p)$, which are used in geometry and mathematical physics, more precisely they include:
(1) Convex elliptic polynomials (which for instance represent the symbols of elliptic differential operators widely studied both in $\mathbb{R}^d$ and on Riemannian manifolds); the most well-studied particular case is surely given by quadratic polynomials, the corresponding Hamiltonians describing the energy of classical mechanical systems in $\mathbb{R}^d$ or on Riemannian manifolds.
(2) Hamiltonians provided by the Lévy-Khinchine formula, namely
$$H(x,p) = \frac{1}{2}\,(G(x)p,p) - (A(x),p) + \int_{\mathbb{R}^d\setminus\{0\}} \Big(e^{-(p,\xi)} - 1 + \frac{(p,\xi)}{1+\xi^2}\Big)\, d\nu_x(\xi),$$
where all elements of $G(x) \in \mathbb{R}^{d\times d}$, $A(x) \in \mathbb{R}^d$ and their derivatives up to order three are bounded, $\mu_1 E_d < G(x) < \mu_2 E_d$ for some $\mu_1, \mu_2 > 0$ and ν_x is a Lévy measure on $\mathbb{R}^d \setminus \{0\}$ with bounded support and some mild regularity assumptions; these Hamiltonians represent the symbols of pseudo-differential operators, which describe generators of most general Feller processes; the solutions of boundary value problem for these Hamiltonians describe the quasi-classical (or small diffusion) asymptotics of trajectories of this processes.

It is well known (see [3], vol. 1, Proposition 2, p.330) that extremal $x(\cdot)$ connecting x_0 and x is a strong minimizer, if it can be embedded into Meyer field covering domain $\Gamma \subset \mathbb{R}^{d+1}$ and
$$\mathcal{E}_L(z, \mathcal{P}(\tau, z), q) > 0 \tag{2.4}$$

for $(t,z) \in \Gamma$ and $q \neq \mathcal{P}(t,z)$, where

$$\mathcal{E}_L(x,p,q) := L(x,q) - L(x,p) - (q-p)\cdot L_p(x,p)$$

is the Weierstrass excess function of the Lagrangian $L(x,p)$ and $\mathcal{P}(t,x)$ is the slope function of Meyer field.

Theorem 2.1 implies the existence of the Meyer field in $\Gamma(x_0) = (0,T)\times B_r(x_0)$ for any $x_0 \in \mathbb{R}^d$ with slope function $\mathcal{P}(t,x) = P(t,x_0,p_0(t,x_0,x))$. Moreover, condition (i) Theorem 2.1 infer that $H_{pp}(x,p) \geq \varepsilon E_d$, which in turn yields (2.4). Hence

$$S(t,x,x_0) := \min I_y(t,x,x_0) = I_{\widetilde{x}(\cdot)}(t,x,x_0)\,, \tag{2.5}$$

where minimum is taken over all y lying completely in $B_r(x_0)$ with the boundary conditions $y(0) = x_0$ and $y(t) = x$.

Lemma 2.2. *Under the conditions of Theorem 2.1 there exist $r_1 \in (0,r]$, $T_1 \in (0,T]$ (r,T as in Theorem 2.1) such that*

$$\max_{x,x_0\in\mathbb{R}^d:|x-x_0|\leq r_1} S(t,x,x_0) < \min_{0<\tau\leq t}\ \min_{x,x_0\in\mathbb{R}^d:|x-x_0|=r} S(\tau,x,x_0) \tag{2.6}$$

for all $0<t<T_1$. Moreover

$$\lim_{t\to 0+}\ \min_{x,x_0\in\mathbb{R}^d:|x-x_0|=r} S(t,x,x_0) = +\infty\,. \tag{2.7}$$

The proofs of Theorem 2.1 and Lemma 2.2 will be given in Sections 4 and 5 respectively.

Observe that (i) Theorem 2.1 implies that $H(x,p)$ is bounded from below and so is $L(x,v)$. Note that functionals (1.2) with Lagrangians $L(x,v)$ and $L(x,v)+c$ with $c\in\mathbb{R}$ being some constant have the same minimizing functions. Therefore without loss of generality we assume that $L(x,v)\geq 0$ and then

$$I_y(t,x,x_0) \geq 0$$

for all piecewise smooth $y(\tau)$.

Corollary 2.3. *Let $0<t<T_1$, $x,x_0\in\mathbb{R}^d$, $|x-x_0|\leq r_1$ (T_1, r_1 as in Lemma 2.2). Then under the conditions of Lemma 2.2 the characteristic $\widetilde{x}(\tau)$ given by* (2.2) *provides the unique absolute minimum for the functional $I_y(t,x,x_0)$.*

Proof. As we have seen above $\widetilde{x}(\tau)$ provides a minimum for $I_y(t,x,x_0)$ among all curves lying completely in $B_r(x_0)$. Let us suppose that $\widetilde{y}(\tau)$ also provides a minimum, and $|\widetilde{y}(s')-x_0| = r$ for some $s'<t$, $r'\leq r$. By (2.5), (2.6) we see

$$I_{\widetilde{y}}(t,x,x_0) \geq I_{\widetilde{y}}(s',\widetilde{y}(s'),x_0) \geq S(s',\widetilde{y}(s'),x_0) > S(t,x,x_0) = I_{\widetilde{x}}(t,x,x_0)\,.$$

□

Theorem 2.4 (Tonelli's theorem)**.** *Under the conditions of Theorem 2.1 for any $t>0$ and $x,x_0\in\mathbb{R}^d$ there exists a characteristic $\widetilde{x}(\tau)$ with boundary conditions $\widetilde{x}(0)=x_0$, $\widetilde{x}(t)=x$ that provides an absolute minimum (probably not unique) for $I_y(t,x,x_0)$ over all piecewise smooth curves $y(\tau)$ connecting x_0 and x in time t.*

Our proof of the implication Lemma 2.2 ⇒ Theorem 2.4 (and also the deduction of Theorem 2.6 given below) is similar to the one given in [4, pp. 56-57]. We give it here for reader's convenience.

Proof of Theorem 2.4. Let $y_n = y_n(\tau)$ will be a minimizing sequence for $I_y(t,x,x_0)$, i.e.

$$I_{y_n}(t,x,x_0) \to \inf_y I_y(t,x,x_0) \tag{2.8}$$

and $y_n(0) = x_0$, $y_n(t) = x$. Let $k \in \mathbb{N}$ be the integer part of $|x - x_0|/r_1$, namely

$$kr_1 \le |x - x_0| < (k+1)r_1 \tag{2.9}$$

and suppose first (in Steps 1 and 2) that

$$t \in (0, T_1) \tag{2.10}$$

(T_1, r, r_1 as in Lemma 2.2).

Step 1. We show that for all $\tau \in (0,t)$ and $n \in \mathbb{N}$

$$y_n(\tau) \in \mathcal{K} \quad \text{for some compact} \quad \mathcal{K} \subset \mathbb{R}^d\,. \tag{2.11}$$

We define $\tau_i = \tau_i(n) \quad i \in \mathbb{N}_0$ by the recurrent formula

$$\tau_0 = 0; \quad \tau_i = \inf\{\tau \in (\tau_{i-1}, t] : |y_n(\tau) - y_n(\tau_{i-1})| = r\}$$

($\inf \emptyset := +\infty$) and put $m = m(n) = \max\{i : \tau_i < \infty\}$. Since $y(0) = y(\tau_0) = x_0$

$$|y_n(\tau) - x_0| \le \sum_{j=1}^{i-1} \big|y_n(\tau_j) - y_n(\tau_{j-1})\big| + \big|y_n(\tau) - y_n(\tau_{i-1})\big| \le ir \le mr$$

for $\tau \in (\tau_{i-1}, \tau_i)$ and

$$|y_n(\tau) - x_0| \le |y_n(\tau) - y_n(\tau_m)| + |y_n(\tau_m) - x_0| \le (m+1)r$$

for $\tau \in (\tau_m, t)$. Thus

$$y_n(\tau) \in B_{(m+1)r}(x_0)\,. \tag{2.12}$$

We show that

$$m(n) \le k+1 \tag{2.13}$$

for all but finite number of n. Then combining (2.12) and (2.13) proves (2.11) with $\mathcal{K} = B_{(k+2)r}(x_0)$.

To prove (2.13) we assume by contradiction that $m \ge k+2$. Let $\widehat{y}_n(\tau)$ be solution of (1.4) on each interval $[\tau_{i-1}, \tau_i]$ $i = 1, \dots, k$ (k as in (2.9)) and on the interval $[\tau_k, t]$ and

$$\widehat{y}_n(\tau_i) = x_0 + \frac{i}{k+1}\,(x - x_0) \quad i = 1, \dots, k\,,$$

$\widehat{y}_n(t) = x$. Since $|\widehat{y}_n(\tau_i) - \widehat{y}_n(\tau_{i-1})| < r_1$ and $|y_n(\tau_i) - y_n(\tau_{i-1})| = r$, an application of (2.6) gives

$$\begin{aligned} S(\tau_i - \tau_{i-1}, y_n(\tau_i), y_n(\tau_{i-1})) &> S(\tau_i - \tau_{i-1}, \widehat{y}_n(\tau_i), \widehat{y}_n(\tau_{i-1})) \\ &= I_{\widehat{y}_n}(\tau_i - \tau_{i-1}, \widehat{y}_n(\tau_i), \widehat{y}_n(\tau_{i-1}))\,. \end{aligned} \tag{2.14}$$

Summing (2.14) over $i = 1, \dots, k$ and using $S(\tau_{k+1} - \tau_k, y_n(\tau_{k+1}), y_n(\tau_k)) \ge S(t - \tau_k, x, \widehat{y}_n(\tau_k))$ we obtain

$$\begin{aligned} I_{y_n}(\tau_{k+1}, y_n(\tau_{k+1}), x_0) &= \sum_{i=1}^{k+1} I_{y_n}(\tau_i - \tau_{i-1}, y_n(\tau_i), y_n(\tau_{i-1})) \\ &\ge \sum_{i=1}^{k+1} S(\tau_i - \tau_{i-1}, y_n(\tau_i), y_n(\tau_{i-1})) \\ &\ge I_{\widehat{y}_n}(t, x, x_0)\,, \end{aligned}$$

and so

$$\inf_n I_{y_n}(\tau_{k+1}, y_n(\tau_{k+1}), x_0) \geq \inf_n I_{\widehat{y}_n}(t, x, x_0)\,.$$

On the other hand, by (2.7) we have

$$\inf_n I_{y_n}(\tau_{k+2} - \tau_{k+1}, y_n(\tau_{k+2}), y_n(\tau_{k+1})) > 0\,.$$

Thus

$$\inf_n I_{y_n}(t, x, x_0) > \inf_n I_{\widehat{y}_n}(t, x, x_0)\,,$$

which contradicts (2.8). This contradiction proves (2.13). Without loss of generality we may assume that (2.13) holds for all n.

Step 2. Since the sequence $(\tau_i, y_n(\tau_i))$ lies in the compact $[0,t] \times \mathcal{K}$ for any $i = 1, \dots, m$, then without loss of generality we assume that

$$\tau_i(n) \to s_i\,, \quad y_n(\tau_i) \to b_i \quad \text{as} \quad n \to \infty$$

$i = 1, \dots, m$ for some $s_i \in (0,t)$, $b_i \in \mathbb{R}^d$. Let $\widetilde{y}_n(\tau)$ be solution of (1.4) on each interval $[\tau_{i-1}, \tau_i]$ $\ i = 1, \dots, m$ and on the interval $[\tau_m, t]$ and

$$\widetilde{y}_n(\tau_i) = y_n(\tau_i) \quad i = 1, \dots, k\,,$$

$\widetilde{y}_n(t) = x$. Clearly

$$I_{y_n}(t, x, x_0) \geq I_{\widetilde{y}_n}(t, x, x_0)\,. \tag{2.15}$$

Let $y(\tau)$ be solution of (1.4) on each interval $[s_{i-1}, s_i]$ $\ i = 1, \dots, m$ and on the interval $[s_m, t]$ and $y(s_i) = b_i$, $i = 1, \dots, m$, $y(t) = x$. Theorem 2.1 implies

$$\lim_{n\to\infty} \|y - \widetilde{y}_n\|_{C^1([s_{i-1}+\delta,\, s_i-\delta])} = 0 \tag{2.16}$$

for any $\delta < (s_{i-1} - s_i)/2$, $i = 1, \dots, m$. Note that for all but finite number i we have $[s_{i-1} + \delta, s_i - \delta] \subset [\tau_{i-1}, \tau_i]$. Therefore,

$$\lim_{n\to\infty} I_{\widetilde{y}_n}(t, x, x_0) = I_y(t, x, x_0)\,. \tag{2.17}$$

Combining (2.15) and (2.17) we get

$$\lim_{n\to\infty} I_{y_n}(t, x, x_0) \geq I_y(t, x, x_0)\,.$$

It suffices to show smoothness of $y(\tau)$ at $\tau = s_i$ $i = 1, \dots, n$. For each $i = 1, \dots, m$ we take $s_i' < s_i < s_i''$ such that $|y(\tau) - y(s_i)| \leq r_1$ for all $\tau \in [s_i', s_i'']$. Since $y(\tau)$ provides the unique minimum among all curves lying completely in $B_{r_1}(y(s_i))$ it should coincide with the solution of (1.4) which passes through points $y(s_i')$, $y(s_i'')$ in times s_i' and s_i'' respectively. In particular this means that y is continuously differentiable, i.e. is a solution of (1.4).

Step 3. Previously we have imposed restriction (2.10). Now let

$$\ell\,\frac{T_1}{2} \leq t < (\ell+1)\frac{T_1}{2}$$

for some $\ell \in \mathbb{N}$ and let $y_n = y_n(\tau)$ be a minimizing sequence. We take $\widetilde{y}_n = \widetilde{y}_n(\tau)$ connecting x_0 and x such that $\widetilde{y}_n$ is characteristic of the system (1.4) on each interval $(kT_1/2, (k+1)T_1/2)$ and $\widetilde{y}_n(kT_1/2) = y_n(kT_1/2)$ for $k = 1, \dots, \ell$.

By above proof we have that $I_{y_n}(t, x, x_0) \geq I_{\widetilde{y}_n}(t, x, x_0)$. As in Step 1 we show that graphs of all $\widetilde{y}_n$ lie in a compact. Hence we may assume that there exist limits $b_k := \lim_{n\to\infty} \widetilde{y}_n(kT_1/2)$, $k = 1, \dots, \ell$. We take $y = y(\tau)$ connecting x_0 and x such that y is characteristic of the system (1.4) on each interval $(kT_1/2, (k+1)T_1/2)$ and

$y(kT_1/2) = b_k$, $k = 1, \dots, \ell$. As in Step 2 we show that $y(\tau)$ is a minimizer and it is smooth at $\tau = kT_1/2$, $k = 1, \dots, \ell$. □

To any solution (X, P) of (1.4) there corresponds the equation in variations

$$\begin{aligned} \dot{v} &= H_{xp}(X(\tau), P(\tau))v + H_{pp}(X(\tau), P(\tau))w \\ \dot{w} &= -H_{xx}(X(\tau), P(\tau))v - H_{xp}(X(\tau), P(\tau))w \end{aligned} \tag{2.18}$$

We say that points $X(\tau_1)$, $X(\tau_2)$ are conjugate if there exists nontrivial (not vanishing identically) solution $(v(\tau), w(\tau))$ of (2.18) such that $v(\tau_1) = v(\tau_2) = 0$.

Proposition 2.5. *Suppose Hamiltonian $H(x, p)$ is smooth and strictly convex in p. If a characteristic $X(\tau)$ contains two conjugate points $X(\tau_1)$, $X(\tau_2)$ then for any $\delta > 0$ the curve $X(\tau)$ $\tau \in [\tau_1, \tau_2 + \delta]$ does not provide even a local minimum among curves joining $X(\tau_1)$ and $X(\tau_2 + \delta)$ in time $\tau_2 - \tau_1 + \delta$.*

The proof of the above Proposition can be found in [1, p.77 Theorem 1]. Following [4] we say that for some $x_0 \in \mathbb{R}^d$ the point $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ is *regular* and write $(t, x) \in \mathrm{Reg}(x_0)$ if (**i**) points x, x_0 are not conjugate; (**ii**) the absolute minimum is attained on a unique curve; (**iii**) this curve is a characteristic $\widetilde{x}(\tau, t, x, x_0)$ with boundary conditions $\widetilde{x}(0) = x_0$, $\widetilde{x}(t) = x$.

Theorem 2.6. *Under the conditions of Theorem 2.1 for any $x_0 \in \mathbb{R}^d$ the set of regular points is open, connected and everywhere dense in $\mathbb{R}_+ \times \mathbb{R}^d$. For any fixed $t > 0$ the set $\{x \in \mathbb{R}^d : (t, x) \in \mathrm{Reg}(x_0)\}$ is open and everywhere dense in $\mathbb{R}^d$.*

Proof. For any $t > 0$, $x, x_0 \in \mathbb{R}^d$ we take characteristic $\widetilde{x}(\tau) = \widetilde{x}(\tau; t, x, x_0)$ which delivers absolute minimum for $I_y(\tau, x, x_0)$. Note that the existence of such $\widetilde{x}(\tau)$ is guaranteed by Theorem 2.4. We show that for any $\tau \in (0, t)$ characteristic $\widetilde{x}(s)$, $s \in [0, \tau]$ delivers the unique absolute minimum to $I_y(\tau, \widetilde{x}(\tau), x_0)$. Clearly $\widetilde{x}(s)$ provides an absolute (a priori not unique) minimum to $I_y(\tau, \widetilde{x}(\tau), x_0)$. If another function $\widetilde{y} = \widetilde{y}(s)$, $s \in [0, \tau]$ with $\widetilde{y}(0) = x_0$, $\widetilde{y}(\tau) = \widetilde{x}(\tau)$ also provides an absolute minimum, then defining it on $(\tau, t]$ by $\widetilde{y}(s) = \widetilde{x}(s)$, $s \in [\tau, t]$ we deduce that $\widetilde{y}$ also provides a minimum for $I_y(t, x, x_0)$. Take $s' < \tau < s''$ such that $\widetilde{y}(s'), \widetilde{y}(s'') \in B_{2^{-1}r_1}(\widetilde{y}(\tau))$, and so $|\widetilde{y}(s') - \widetilde{y}(s'')| \le r_1$. By Corollary 2.3 the characteristic connecting $\widetilde{y}(s')$ and $\widetilde{y}(s'')$ in time $s'' - s'$ provides the unique absolute minimum. Therefore $\widetilde{x}(s) = \widetilde{y}(s)$ for $s \in [s', s'']$. Consequently we have $\widetilde{x}(s) = \widetilde{y}(s)$ for all $s \in [0, \tau]$.

By Proposition 2.5 $\widetilde{x}(0)$ and $\widetilde{x}(\tau)$ are not conjugate. Hence $(\tau, \widetilde{x}(\tau))$ is regular. This immediately implies that Reg (x_0) is connected and everywhere dense in $\mathbb{R}_+ \times \mathbb{R}^d$. Moreover, x-projection of Reg (x_0) is everywhere dense in $\mathbb{R}^d$. To prove that Reg (x_0) is open, note that if (t, x) is regular then $\partial X(t, x_0, p_0)/\partial p_0 \neq 0$. An application of implicit function theorem gives that $p_0 = p_0(t, x, x_0)$ is well-defined by (2.3) in some neighborhood $U(t, x) \subset \mathbb{R}^{d+1}$ of (t, x). As we know the characteristic $\widetilde{x}(\tau) = \widetilde{x}(\tau; t', x', x_0)$ delivers local minimum for any $(t', x') \in U(t, x)$. Since the quantity $\min_y I_y(t', x', x_0)$ depends continuously on (t', x') and for $(t', x') = (t, x)$ the unique absolute minimum is provided by $\widetilde{x}(\tau; t, x, x_0)$, then $x(\tau; t', x', x_0)$ provides an absolute minimum for any $(t', x') \in U_0(t, x)$ for some $U_0(t, x) \subset U(t, x)$.

Similarly fixing $t > 0$ one shows that the set $\{x : (t, x) \in \mathrm{Reg}(x_0)\}$ is open. □

3. Auxiliary results

In this section we will deduce some implications from (2.1) and (i), (ii) Theorem 2.1.

Notation. Throughout this paper $\varepsilon > 0$ will be from the conditions of Theorem 2.1 and $M = M(h) > 0$ will be the constant from (2.1). In sections 3 - 5 we will construct constants $M_1, \dots, M_{11} > 0$, $\rho > 0$,

$$1 > t_1 \ge t_2 \ge T \ge T_1 > 0\,, \quad 1 > c_1 \ge c_2 > 0\,, \quad r \ge r_1 > 0$$

which depend only on ε and $h(z)$.

Lemma 3.1. *The functions $h \in \mathcal{F}$ satisfy the following inequalities*

(i) $|h^{(n)}(z)(h(z))^{n-1}| \le |h'(z)|^n$ $n \in \mathbb{N}_0$,

(ii) $\Big|h^{(n)}\Big(z + \lambda \frac{|h(z)|}{|h'(z)|}\Big)\Big| \le 3|h^{(n)}(z)|$, $n \in \mathbb{N}_0$

for all $z > M$ and $|\lambda| \le 1$.

Proof. The case $n = 0$ in (i) is trivial. We assume that $n \ge 1$. By repeated application of (2.1) we obtain

$$|h^{(n)}(z)(h(z))^{n-1}| \le |h^{(n-1)}(z)(h(z))^{n-2}h'(z)| \le \dots \le |h'(z)|^n$$

and (i) follows. We proceed with (ii). Using Taylor's development $h^{(n)}(z + a) = \sum_{m=0}^{\infty} h^{(n+m)}(z)a^m/m!$ with $a = |\lambda h(z)/h'(z)|$ gives

$$h^{(n)}\Big(z + \lambda \frac{|h(z)|}{|h'(z)|}\Big) = \sum_{m=0}^{\infty} |\lambda|^m \frac{h^{(m+n)}(z)}{m!} \frac{|h(z)|^m}{|h'(z)|^m}\,.$$

Since, by (2.1),

$$|h^{(m+n)}(z)(h(z))^m| \le |h^{(m+n-1)}(z)h'(z)(h(z))^{m-1}| \le \dots \le |h^{(n)}(z)||h'(z)|^m\,,$$

it follows

$$\Big|h^{(n)}\Big(z + \lambda \frac{|h(z)|}{|h'(z)|}\Big)\Big| \le |h^{(n)}(z)| \sum_{m=0}^{\infty} \frac{|\lambda|^m}{m!} \le |h^{(n)}(z)| \exp\{1\},$$

where we used $|\lambda| < 1$. □

Lemma 3.2. *Let $h \in \mathcal{F}$, $H \in C^4(\mathbb{R}^{2d})$ be such that conditions (i), (ii) of Theorem 2.1 hold. Then for some constants $M_1, M_2, M_3 > 0$*

(i) *$|h^{(m)}(z)|$ is monotone on $(M_1, +\infty)$ for all $m = 0, \dots, 3$,*

(ii) *$|h'(z)|$ increases on $(M_1, +\infty)$,*

(iii) *$|H_p(x, p)| \ge \varepsilon |h'(|p|)| - M_2$ for all $x, p \in \mathbb{R}^d$,*

(iv) *$(p, H_p(x, p)) \ge \varepsilon |h'(|p|)|\,|p| - M_3$ for all $x, p \in \mathbb{R}^d$.*

Proof. An application of (2.1) with $n = m + 2$, $m \in \mathbb{N}_0$ implies

$$\big((g(z))^2\big)'' = 2g''(z)g(z) + 2(g'(z))^2 > 0$$

for $g(z) = h^{(m)}(z)$. Therefore, $g(z)$ has at most two intervals of monotonicity. Hence $g(z)$ is monotone for $z \ge M_{1,m}$ for some $M_{1,m} > 0$. Take

$$M_1 = \varepsilon^{-1} + \max_{m=0,\dots,3} M_{1,m}$$

and (i) follows.

By (i) Theorem 2.1 we have $H_{pp}(x,p) \geq \varepsilon E_d$. Hence

$$|p|\,|H_p(x,p)| \geq \big(p, H_p(x,p)\big) = \big(p, H_p(x,0)\big) + \int_0^1 \big(p, H_{pp}(x,\tau p)p\big)\,d\tau \tag{3.1}$$

$$\geq \big(p, H_p(x,0)\big) + \varepsilon\,|p|^2 \tag{3.2}$$

$$\geq \varepsilon\,|p|^2 - \varepsilon^{-1}|p|\,, \tag{3.3}$$

where we used that, by (ii) Theorem 2.1, $|H_p(x,0)| \leq \varepsilon^{-1}$. Hence

$$|H_p(x,p)| \geq \varepsilon\,|p| - \varepsilon^{-1}\,.$$

Using (ii) Theorem 2.1 we get

$$|h'(z)| \geq \varepsilon\,z - \varepsilon^{-1} \tag{3.4}$$

for $z \geq M_1$ (recall $M_1 \geq \varepsilon^{-1}$). This and (i) Lemma 3.2 imply (ii) Lemma 3.2.

We proceed with (iii) and (iv). Without loss of generality we may assume that $h'(z)$ increases for $z \geq M_1$. (Otherwise we replace h by $-h$ noticing that the conditions of Theorem 2.1 for $-h$ and H still hold.) So

$$h''(z) > 0 \quad \text{for all} \quad z \geq M_1\,.$$

Hence for $|p| \geq M_1$ we get

$$\int_0^1 |h''(\tau|p|)|\,|p|\,d\tau \geq \int_{\frac{M_1}{|p|}}^1 h''(\tau|p|)\,|p|\,d\tau = h'(|p|) - h'(M_1)\,.$$

Consequently, due to (i) Theorem 2.1,

$$\int_0^1 \big(p, H_{pp}(x,\tau p)p\big)\,d\tau \geq \varepsilon|p|^2 \int_0^1 (1 + |h''(\tau|p|)|)\,d\tau$$

$$\geq \varepsilon|p|\,(|p| + h'(|p|) - h'(M_1))\ .$$

Using this and the first line in (3.1) we have

$$(p, H_p(x,p)) \geq -\varepsilon^{-1}|p| + \varepsilon|p|\,(|p| + h'(|p|) - h'(M_1)) \tag{3.5}$$

$$\geq \varepsilon h'(|p|)\,|p| - M_3\,, \tag{3.6}$$

where

$$M_3 := -\min_{z>0}\{-\varepsilon^{-1}z + \varepsilon z(z - h'(M_1))\}\,.$$

Using again (3.5) we get

$$|H_p(x,p)| \geq -\varepsilon^{-1} + \varepsilon\,(|p| + h'(|p|) - h'(M_1)) \geq \varepsilon h'(|p|) - \varepsilon^{-1} - \varepsilon h'(M_1)\,, \tag{3.7}$$

where $M_2 := \varepsilon^{-1} + \varepsilon h'(M_1)$. □

Lemma 3.3. *Under the conditions of Theorem 2.1 there exist $t_1, M_4 > 0$ such that for all $c \in (0,1/2)$, $t \in (0,t_1)$, $x_0 \in \mathbb{R}^d$,*

$$p_0 \in V_{c,t} := \{p \in \mathbb{R}^d : t|h'(|p|)| < c\} \tag{3.8}$$

and $\tau \in [0,t]$ we have

$$|h^{(n)}(|P(\tau)|)| \leq M_4(|h^{(n)}(|p_0|)| + 1) \quad n = 0,1,2. \tag{3.9}$$

Moreover,

$$|X(t,x_0,p_0) - x_0| \leq M_4 t(|h'(|p_0|)| + 1)\,, \tag{3.10}$$

$$|P(t,x_0,p_0) - p_0| \leq M_4 t(|h(|p_0|)| + 1)\,. \tag{3.11}$$

Proof. We choose $0 < t_1 < 1$ such that the system (1.4) has a solution on $[-t_1, t_1]$. Let us first (in Steps 1 and 2) assume that

$$|P(\tau)| \geq M_1 \quad \text{for all} \quad \tau \in [0, t_1]\,. \tag{3.12}$$

Step 1. Notice that (i), (ii) Lemma 3.2 imply that

$$|h(z)| \quad \text{increases for} \quad z \geq M_1\,. \tag{3.13}$$

In particular $h(z)$ preserves sign for $z \geq M_1$. Without loss of generality we may assume (in Steps 1 and 2) that $h(z) > 0$ for $z \geq M_1$. Using (ii) Theorem 2.1 and $M_1 > \varepsilon^{-1}$ we find from (3.12) and the system (1.4)

$$|P(t) - p_0| \leq \int_0^t |H_x(X(\tau), P(\tau))|\, d\tau \leq \int_0^t h(|P(\tau)|)\, d\tau$$

and so

$$|p_0| - \int_0^t h(|P(\tau)|)\, d\tau \leq |P(t)| \leq |p_0| + \int_0^t h(|P(\tau)|)\, d\tau\,.$$

Due to (3.13) and Gronwall Lemma,

$$y_1(\tau) \leq |P(\tau)| \leq y_2(\tau),$$

where $y_i(\tau)$ $i = 1, 2$ solve the equation

$$\dot{y}_i(\tau) = (-1)^i\, h(y_i(\tau))\,, \quad y_i(0) = |p_0|.$$

One readily sees that $y_i(\tau) = \Phi^{-1}\left(\Phi(|p_0|) + (-1)^i \tau\right)$, $i = 1, 2$, where

$$\Phi(\lambda) = \int_{M_1}^{\lambda} \frac{d\zeta}{h(\zeta)}\,.$$

Hence

$$\Phi^{-1}\left(\Phi(|p_0|) - \tau\right) \leq |P(\tau)| \leq \Phi^{-1}\left(\Phi(|p_0|) + \tau\right). \tag{3.14}$$

Step 2. An application of Taylor's formula on Φ^{-1} yields

$$\Phi^{-1}\left(\Phi(z) + \tau\right) = z + \sum_{n=1}^{\infty} \frac{a_n(z)}{n!}\, \tau^n, \tag{3.15}$$

where

$$a_n(z) = \frac{d^n}{d\tau^n}\, \Phi^{-1}\left(\Phi(z) + \tau\right)\big|_{\tau=0} = (\Phi^{-1})^{(n)}(\Phi(z))$$

or

$$a_n(\Phi^{-1}(\kappa)) = (\Phi^{-1})^{(n)}(\kappa) \tag{3.16}$$

with $\kappa = \Phi(z)$. We differentiate (3.16) in κ to get

$$a_n'(\Phi^{-1}(\kappa))(\Phi^{-1}(\kappa))' = (\Phi^{-1})^{(n+1)}(\kappa) = a_{n+1}(\Phi^{-1}(\kappa))\,. \tag{3.17}$$

Consequently, (3.17) and $(\Phi^{-1}(\kappa))' = dz/d\kappa = (\Phi'(z))^{-1} = h(z)$ imply

$$a_1(z) = h(z) \quad \text{and} \quad a_{n+1}(z) = a_n'(z) h(z)\,. \tag{3.18}$$

Using induction we deduce from (3.18)

$$a_n(z) = {\sum}' h^{(m_1)}(z) \dots h^{(m_{n-1})}(z) h(z) \quad n \geq 2\,, \tag{3.19}$$

where the sum $\sum'$ is taken over some $m_1, \dots, m_{n-1} \in \mathbb{N}_0$ such that $m_1 + \dots + m_{n-1} = n - 1$ and contains $(n-1)!$ terms. It follows

$$|a_n(z)| \leq (n-1)! \max_{m_1 + \dots + m_{n-1} = n-1} \left| h^{(m_1)}(z) \dots h^{(m_{n-1})}(z) h(z) \right|.$$

By (i) Lemma 3.1 and because of $\sum_{k=1}^{n-1}(m_k-1)=0$ we have

$$|\, h^{(m_1)}(z)\dots h^{(m_{n-1})}(z)| = \prod_{k=1}^{n-1} |\, h^{(m_k)}(z)(h(z))^{m_k-1}| \le |h'(z)|^{n-1}.$$

Consequently,

$$|a_n(z)\tau^n| \le (n-1)!\, |\tau h'(z)|^{n-1}\, \tau |h(z)|.$$

Since, by (3.8), $|\tau h'(|p_0|)| \le c$, it follows that

$$|\frac{a_n(|p_0|)\tau^n}{n!}| < c^{n-1}\, \tau |h(|p_0|)|. \tag{3.20}$$

Substituting (3.20) in (3.15) we get

$$\Phi^{-1}\left(\Phi(|p_0|)+\tau\right) \le |p_0| + \frac{1}{1-c}\, \tau|h(|p_0|)| \le |p_0| + 2\tau|h(|p_0|)| \tag{3.21}$$

for $c\in(0,1/2)$. Similarly we obtain

$$\Phi^{-1}\left(\Phi(|p_0|)-\tau\right) \ge |p_0| - 2\tau|h(|p_0|)| \tag{3.22}$$

for $c\in(0,1/2)$. Combining (3.14), (3.21) and (3.22) we obtain

$$|p_0| - 2\tau|h(|p_0|)| \le |P(\tau)| \le |p_0| + 2\tau|h(|p_0|)|$$

for $c\in(0,1/2)$. Using $p_0\in V_{c,t}$ gives $\tau\le t\le c/|h'(|p_0|)|$ and so

$$|p_0| - 2c\,\frac{|h(|p_0|)|}{|h'(|p_0|)|} \le |P(\tau)| \le |p_0| + 2c\,\frac{|h(|p_0|)|}{|h'(|p_0|)|}. \tag{3.23}$$

An application of (ii) Lemma 3.1 with $\lambda = 2c \le 1$ gives

$$\begin{aligned} |h^{(n)}(|P(\tau)|)| &\le \max\big\{h\big(|p_0| + 2c\,\frac{|h(|p_0|)|}{|h'(|p_0|)|}\big), h\big(|p_0| - 2c\,\frac{|h(|p_0|)|}{|h'(|p_0|)|}\big)\big\} \\ &\le 3|h^{(n)}(|p_0|)| \quad n=0,1,2. \end{aligned} \tag{3.24}$$

Similarly one can check (we will need this in the proof of Lemma 3.4) that

$$|h^{(n)}(|P(\tau_1)|)| \le 3|h^{(n)}(|P(\tau_2)|)| \quad n=0,1,2 \tag{3.25}$$

for any $\tau_1,\tau_2\in[0,t]$.

Step 3. Finally we suppose that (3.12) does not hold. If $|P(\tau)|\le M_1$ for all $\tau\in[0,t_1]$ then taking

$$M_4 := 3 + 3\max_{z\in[0,\,M_1],n=0,1,2} |h^{(n)}(z)| \tag{3.26}$$

we have

$$|h^{(n)}(|P(\tau)|)| \le \frac{1}{3}\, M_4 \quad n=0,1,2 \tag{3.27}$$

for all τ. If $|P(s_0)| = M_1$ and $|P(\tau)|\ge M_1$ on $[s_0,s_1]\subset[0,t_1]$ for some $s_0<s_1$, then (3.24) is applicable to $P(\tau)$ on $[s_0,s_1]$ and so

$$|h^{(n)}(|P(\tau)|)| \le 3|h^{(n)}(|M_1|)| \le M_4 \quad n=0,1,2 \tag{3.28}$$

for $\tau\in[s_0,s_1]$. Combining (3.27) and (3.28) we have

$$|h^{(n)}(|P(\tau)|)| \le M_4 \quad n=0,1,2. \tag{3.29}$$

Combining this and (3.24) we arrive at (3.9). Estimates (3.10), (3.11) are direct consequences of (ii) Theorem 2.1 and (3.9). □

Lemma 3.4. *Let us define* $\omega : [M_4, +\infty) \to \mathbb{R}$ *by the identity*

$$\omega(|h'(z)|) \equiv z\,. \tag{3.30}$$

Then there exist $c_1 \in (0, 1/2)$ *and constants* $M_5, M_6 > 0$ *such that for all* $t \in (0, t_1)$, $c \in (0, c_1)$, $x_0 \in \mathbb{R}^d$, $p_0 \in V_{c,\,t}$ *one has*

(i)

$$|p_0| \geq \omega\big(M_4^{-1}\,\frac{|x - x_0|}{t} - M_5\big) \quad \textit{provided that } M_4^{-1}\frac{|x - x_0|}{t} - M_5 > M_4\,,$$

(ii)

$$|p_0| \leq \omega\big(\frac{2}{\varepsilon}\,\frac{|x - x_0|}{t} + M_5\big)\,,$$

(iii)

$$\big(p_0, \frac{x - x_0}{t}\big) \geq \frac{\varepsilon}{2}\,|p_0|\big(M_4^{-1}\frac{|x - x_0|}{t} - M_5\big) - M_6\,,$$

where we write for short $x = X(t, x_0, p_0)$.

Remark 3.5. Notice that (ii) Lemma 3.2 and the fact that $M_4 > |h'(M_1)|$ imply the correctness of definition (3.30).

Proof. Step 1. In this step we show that

$$|(p_0, H_p(X(\tau), P(\tau)) - H_p(x_0, p_0))| \leq 2^{-1}\varepsilon\,|h'(|p_0|)||p_0| + 2\omega(M_4)(\varepsilon^{-1} + M_4)\,. \tag{3.31}$$

Let us first assume that

$$|P(\tau)| > M_1 \quad \forall \tau \in [0, t]\,. \tag{3.32}$$

From (1.4), (3.25) and (ii) Theorem 2.1 (recall $M_1 > \varepsilon^{-1}$) we deduce

$$\begin{aligned}\sum_{j=1}^{d} |(X(\tau) - x_0)_j| &\leq \sum_{j=1}^{d} \int_0^t \Big|\frac{\partial H}{\partial p_j}\,(X(\tau), P(\tau))\Big|\,d\tau \\ &\leq t \max_{0 \leq \tau \leq t} |h'(|P(\tau)|)| \leq 3t\big|h'(|p_0|)\big|\end{aligned}$$

and

$$\sum_{j=1}^{d} |(P(\tau) - p_0)_j| \leq t \max_{0 \leq \tau \leq t} |h(|P(\tau)|)| \leq 3t\big|h(|p_0|)\big|\,.$$

Denote by

$$R := H_p(X(\tau), P(\tau)) - H_p(x_0, p_0)\,, \tag{3.33}$$

$R = (R_1, \dots, R_d)$. Using the mean value theorem we get

$$\begin{aligned}R_i &= \sum_{j=1}^{d} H_{p_i x_j}(\nu_i, \xi_i)(X(t) - x_0)_j + \sum_{j=1}^{d} H_{p_i p_j}(\nu_i, \xi_i)(P(t) - p_0)_j \\ &=: R_i^1 + R_i^2\end{aligned}$$

for some $\nu_i \in [x_0, X(\tau)]$, $\xi_i \in [p_0, P(\tau)]$, $\quad i = 1, \dots, d$. It follows that

$$|R_i^2| \leq \max_{i,j=1,\dots,d} \big|H_{p_i p_j}(\nu_i, \xi_i)\big|\,\Big(\sum_{j=1}^{d} |(P(\tau) - p_0)_j|\Big)\,.$$

Due to $|\xi_i| > M_1 > \varepsilon^{-1}$ and (ii) Theorem 2.1 we get

$$\big|H_{p_i p_j}(\nu_i, \xi_i)\big| \leq \big|h''(|\xi_i|)\big| \leq \max\{|h''(|p_0|)|, |h''(|P(\tau)|)|\} \leq 3\big|h''(|p_0|)\big|$$

$i,j=1,\dots,d$ and so

$$|R_i^2|\le 9t\big|h''(|p_0|)h(|p_0|)\big|\le 9t\big|h'(|p_0|)\big|^2\,.$$

Similarly $|R_i^1|\le 9t\big|h'(|p_0|)\big|^2$. Consequently,

$$|R|\le\sum_{i=1}^d|R_i|\le 18\,dt\big|h'(|p_0|)\big|^2\,.$$

Setting $c_1=\varepsilon/(36\,d)$ and using $t\le|h'(|p_0|)|^{-1}c$ we get (under assumption (3.32))

$$|H_p(X(\tau),P(\tau))-H_p(x_0,p_0)|\le\frac{\varepsilon}{2}\,|h'(|p_0|)|\,. \tag{3.34}$$

If (3.32) does not hold then, by (3.29) and (ii) Theorem 2.1,

$$|H_p(X(\tau),P(\tau))|\le\varepsilon^{-1}+|h'(|P(\tau)|)|\le\varepsilon^{-1}+M_4\,.$$

Using, by (3.29), $|h'(|p_0|)|\le M_4$ and the fact that $z\ge\omega(M_4)$ implies $|h'(z)|>M_4$ we obtain $|p_0|\le\omega(M_4)$. Hence

$$|(p_0,H_p(X(\tau),P(\tau))-H_p(x_0,p_0))|\le 2\omega(M_4)(\varepsilon^{-1}+M_4)\,.$$

Combining this and (3.26) we arrive at (3.31).

Step 2. If (3.32) holds then, by (3.34)

$$\begin{aligned}|x-x_0|-t|H_p(x_0,p_0)|&\ge-|x-x_0-tH_p(x_0,p_0)|\\&=-\Big|\int_0^t(H_p(X(\tau),P(\tau))-H_p(x_0,p_0))\,d\tau\Big|\\&\ge-\frac{t\varepsilon}{2}\,|h'(|p_0|)|\,,\end{aligned}$$

and so using (iii) in Lemma 3.2, we get

$$|x-x_0|\ge\frac{\varepsilon t}{2}\,|h'(|p_0|)|-tM_2\,. \tag{3.35}$$

If (3.32) does not hold, then, due to (3.29),

$$|x-x_0|\ge 0\ge t|h'(|p_0|)|-M_4t\,. \tag{3.36}$$

Combining (3.35) and (3.36) (recall $\varepsilon<1$) we arrive at

$$\frac{|x-x_0|}{t}\ge\frac{\varepsilon}{2}\,|h'(|p_0|)|-M_2-M_4\,.$$

This and (3.10) yield

$$M_4^{-1}\frac{|x-x_0|}{t}-M_5\le|h'(|p_0|)|\le\frac{2}{\varepsilon}\,\frac{|x-x_0|}{t}+M_5\,, \tag{3.37}$$

where $M_5:=2\varepsilon^{-1}(M_2+M_4)$. Applying $\omega(\cdot)$ to both sides of (3.37) proves (i) and (ii) Lemma 3.4.

Step 3. An application of (3.31) and (iv) Lemma 3.2 give

$$\begin{aligned}(p_0,x-x_0)&=t\,(p_0,H_p(x_0,p_0))+\int_0^t\big(p_0,H_p(X(\tau),P(\tau))-H_p(x_0,p_0)\big)\,d\tau &(3.38)\\&\ge t\,(p_0,H(x_0,p_0))-\frac{\varepsilon t}{2}\,|h'(|p_0|)|\,|p_0|-2\omega(M_4)(\varepsilon^{-1}+M_4)t &(3.39)\\&\ge\frac{\varepsilon t}{2}\,|h'(|p_0|)|\,|p_0|-M_3t-2\omega(M_4)(\varepsilon^{-1}+M_4)t\,. &(3.40)\end{aligned}$$

Using the first inequality in (3.37) and setting $M_6 := M_3 + 2\omega(M_4)(\varepsilon^{-1} + M_4)$ we complete the proof. □

4. Proof of Theorem 2.1

Let

$$\|B\| := \max_{1\le i\le n,\, 1\le j\le m} |(B)_{ij}| \tag{4.1}$$

for any $B \in \mathbb{R}^{n\times m}$. We say that $R(t) = O(r(t)) \quad t \in \mathbf{T} \subset \mathbb{R}$, where $R(t) \in \mathbb{R}^{n\times m}$, $r(t) \in \mathbb{R}_+$ if

$$\|R(t)\| \le Cr(t) \tag{4.2}$$

for some constant $C > 0$ and all $t \in \mathbf{T}$.

Lemma 4.1. *Let c_1, t_1 be as in Lemma 3.4 and the conditions of Theorem 2.1 be satisfied. Then for all $t \in (0, t_1)$, $c \in (0, c_1)$, $x_0 \in \mathbb{R}^d$, $p_0 \in V_{c,t}$ one has*

$$\frac{\partial X(s)}{\partial p_0} = sH_{pp}(x_0, p_0) + s\gamma O(c + t), \tag{4.3}$$

$$\frac{\partial P(s)}{\partial p_0} = E_d + O(c + t), \tag{4.4}$$

where $s \in [0, t]$, E_d is the identity matrix,

$$\gamma = |h''(|p_0|)| + 1 \tag{4.5}$$

and $O(\cdot)$ is taken uniformly in x_0, p_0 and s.

Proof. Step 1. From the system (1.4) we find

$$\ddot X(\tau) = A(X(\tau), P(\tau)), \quad A(x, p) = H_{px}H_p - H_{pp}H_x. \tag{4.6}$$

Here we omit x, p in arguments of the derivatives of H, the terms $H_{px}H_p$, $H_{pp}H_x$ are understood as matrix products and $A(x, p) \in \mathbb{R}^d$. Differentiating with respect to p_0 the Taylor's developments

$$\begin{aligned} X(s) &= x_0 + \dot X(0)s + \int_0^s (s - \tau)\, \ddot X(\tau)\, d\tau, \\ P(s) &= p_0 + \int_0^s \dot P(\tau)\, d\tau \end{aligned} \tag{4.7}$$

and using $\partial \dot X(0)/\partial p_0 = H_{pp}(x_0, p_0)$ gives

$$\frac{\partial X(s)}{\partial p_0} = H_{pp}(x_0, p_0)s + \int_0^s (s - \tau)\, \frac{\partial \ddot X(\tau)}{\partial p_0}\, d\tau \tag{4.8}$$

and

$$\frac{\partial P(s)}{\partial p_0} = E_d + \int_0^s \frac{\partial \dot P(\tau)}{\partial p_0}\, d\tau \tag{4.9}$$

respectively, where

$$\frac{\partial \ddot X(\tau)}{\partial p_0} = A_x(X(\tau), P(\tau))\, \frac{\partial X(\tau)}{\partial p_0} + A_p(X(\tau), P(\tau))\, \frac{\partial P(\tau)}{\partial p_0}\,, \tag{4.10}$$

$$\frac{\partial \dot P(\tau)}{\partial p_0} = -H_{xx}(X(\tau), P(\tau))\frac{\partial X(\tau)}{\partial p_0} - H_{xp}(X(\tau), P(\tau))\frac{\partial P(\tau)}{\partial p_0}\,. \tag{4.11}$$

We take $2d \times 2d$ matrix

$$\ell(s,\tau) = \begin{pmatrix} s^{-1}(s-\tau)\tau A_x(X(\tau),P(\tau)) & s^{-1}(s-\tau)\gamma^{-1}A_p(X(\tau),P(\tau)) \\ -\tau\gamma H_{xx}(X(\tau),P(\tau)) & -H_{xp}(X(\tau),P(\tau)) \end{pmatrix},$$

where γ is given by (4.5). Define a mapping $L : C([0,t],\mathbb{R}^{2d\times d}) \to C([0,t],\mathbb{R}^{2d\times d})$ by

$$[LM](s) := \int_0^s \ell(s,\tau)M(\tau)\,d\tau\,,$$

$M(\cdot) \in C([0,t],\mathbb{R}^{2d\times d})$. We rewrite (4.8)-(4.9) as $\mathcal{M} = \mathcal{M}_0 + L\mathcal{M}$, where

$$\mathcal{M} = \begin{pmatrix} \frac{1}{s}\,\frac{\partial X(s)}{\partial p_0} \\ \gamma\,\frac{\partial P(s)}{\partial p_0} \end{pmatrix} \quad \text{and} \quad \mathcal{M}_0 = \begin{pmatrix} H_{pp}(x_0,p_0) \\ \gamma E_d \end{pmatrix}.$$

Using (ii) in Theorem 2.1, we get $\|\mathcal{M}_0\| = O(\gamma)$. Thus

$$\mathcal{M} = \sum_{n=0}^{\infty} L^n \mathcal{M}_0 = \mathcal{M}_0 + \|L\| O(\gamma) \tag{4.12}$$

provided $\|L\| < 1$.

Step 2. We now estimate $\|L\|$. Using the elementary formula

$$\|B_1B_2\| \le \nu\|B_1\|\|B_2\| \quad \forall B_1 \in \mathbb{R}^{n\times\nu}, B_2 \in \mathbb{R}^{\nu\times m} \tag{4.13}$$

we obtain

$$\max_{0\le s\le t}\Big\|\int_0^s \ell(s,\tau)M(\tau)\,d\tau\Big\| \le sd \max_{0\le\tau\le s\le t}\|\ell(s,\tau)\| \max_{0\le\tau\le t}\|M(\tau)\|\,.$$

Hence

$$\|L\| \le sd \max_{0\le\tau\le s\le t}\|\ell(s,\tau)\|\,.$$

Since $\tau \le s$ and $s^{-1}(s-\tau) \le 1$, it follows that

$$\begin{aligned} \max_{0\le\tau\le s\le t}\|\ell(s,\tau)\| &\le s\max_{0\le\tau\le t}\|A_x(X(\tau),P(\tau))\| + \gamma^{-1}\max_{0\le\tau\le t}\|A_p(X(\tau),P(\tau))\| \\ &\quad + s\gamma\max_{0\le\tau\le t}\|H_{xx}(X(\tau),P(\tau))\| + \max_{0\le\tau\le t}\|H_{xp}(X(\tau),P(\tau))\|\,. \end{aligned} \tag{4.14}$$

Applying (2.1) and (ii) Theorem 2.1 to (4.6), we have

$$\begin{aligned} \|A(x,p)\| &\le d\|H_{xp}\|\,\|H_p\| + d\|H_{pp}\|\,\|H_x\| \\ &\le d|h'(|p|)|^2 + d\,|h''(|p|)h'(|p|)| \\ &\le 2d\,|h'(|p|)|^2, \end{aligned} \tag{4.15}$$

provided $|p| \ge M_1$. Similarly we obtain

$$\begin{aligned} \|A_x(x,p)\| &\le 4d\,|h'(|p|)|^2\,, \\ \|A_p(x,p)\| &\le 3d\big|h''(|p|)\,h'(|p|)\big| + d\big|h'''(|p|)\,h(|p|)\big| \\ &\le 4d\big|h''(|p|)\,h'(|p|)\big| \end{aligned}$$

for $|p| \ge M_1$. Consequently,

$$A_x(x,p) = (h'(|p|))^2 O(1) + O(1), \tag{4.16}$$

$$A_p(x,p) = \big|h''(|p|)\,h'(|p|)\big| O(1) + O(1) \tag{4.17}$$

for all $p \in \mathbb{R}^d$. Hence

$$s^2 A_x(X(\tau), P(\tau)) = O([sh'(|P(\tau)|)]^2) + O(s^2)\,.$$

By (3.9), we get

$$s|h'(|P(\tau)|)| \le M_4(s|h'(|p_0|)| + s) \le M_4(c+s) \tag{4.18}$$

and so

$$s^2 A_x(X(\tau), P(\tau)) = O(c^2 + t^2). \tag{4.19}$$

Since, by (3.9), $\gamma^{-1}|h''(|P(\tau)|)| = O(1)$, then using (4.17), (4.18) and $\gamma^{-1} \le 1$ we find

$$s\gamma^{-1} A_p(X(\tau), P(\tau)) = O(|sh'(P(\tau))|) + \gamma^{-1}O(s) = O(c+t). \tag{4.20}$$

Piecing together (4.19) and (4.20) gives

$$s^2 \max_{0\le\tau\le t} \|A_x(X(\tau), P(\tau))\| + s\gamma^{-1} \max_{0\le\tau\le t} \|A_p(X(\tau), P(\tau))\| = O(c+t)\,. \tag{4.21}$$

Similarly we get

$$s^2\gamma \max_{0\le\tau\le t} \|H_{xx}(X(\tau), P(\tau))\| + s \max_{0\le\tau\le t} \|H_{xp}(X(\tau), P(\tau))\| = O(c+t)\,. \tag{4.22}$$

Combining (4.14), (4.21), (4.22) we arrive at $s \max_{0\le\tau\le s\le t} \|\ell(s,\tau)\| = O(c+t)$. Hence $\|L\| = O(c+t)$. Consequently we rewrite (4.12) as

$$\frac{1}{s}\frac{\partial X(s)}{\partial p_0} = H_{pp}(x_0, p_0) + \gamma O(c+t), \tag{4.23}$$

$$\gamma \frac{\partial P(s)}{\partial p_0} = \gamma(E_d + O(c+t)). \tag{4.24}$$

Multiplying (4.23) and (4.24) by s and γ^{-1} respectively we complete the proof. $\square$

Corollary 4.2. *Under the conditions of Theorem 2.1 there exist $c_2 \in (0, c_1)$, $t_2 \in (0, t_1)$ such that for all $t \in (0, t_2)$, $c \in (0, c_2)$ and $p_1, p_2 \in V_{c,t}$ one has*

$$|X(t, x_0, p_2) - X(t, x_0, p_1)|^2 \ge \frac{t^2\varepsilon^2}{2}\, |p_1 - p_2|^2\,. \tag{4.25}$$

Proof. One has

$$\begin{aligned} &|X(t, x_0, p_2) - X(t, x_0, p_1)|^2 \\ &= \int_0^1 \int_0^1 (p_2 - p_1)^T \Big(\frac{\partial X}{\partial p_0}(t, x_0, p_1 + s(p_2 - p_1))\Big)^T \\ &\quad \times \Big(\frac{\partial X}{\partial p_0}(t, x_0, p_1 + \tau(p_2 - p_1))\Big)\, (p_2 - p_1)\; d\tau\, ds\,, \end{aligned} \tag{4.26}$$

where $(\partial X/\partial p_0)^T$ is transposed matrix to $\partial X/\partial p_0$. We denote by

$$q_1 = p_1 + s(p_2 - p_1)\,, \quad q_2 = p_1 + \tau(p_2 - p_1)\,.$$

From (4.3), we obtain

$$\begin{aligned} \Big(\frac{\partial X(t, x_0, q_1)}{\partial p_0}\Big)^T \frac{\partial X(t, x_0, q_2)}{\partial p_0} &= t^2(H_{pp}(x_0, q_1) + \gamma_1 R_1)(H_{pp}(x_0, q_2) + \gamma_2 R_2) \\ &=: t^2 H_{pp}(x_0, q_1) H_{pp}(x_0, q_2) + t^2 R, \end{aligned} \tag{4.27}$$

where $\gamma_i = |h''(|q_i|)| + \varepsilon^{-1}$, $i = 1, 2$ and $\|R_i\| \le (c+t)\kappa$ for some constant $\kappa > 1$. Using (4.13) we get

$$\|R\| \le d\left(\gamma_1 \|H_{pp}(x_0, q_2)\| \, \|R_1\| + \gamma_2 \|H_{pp}(x_0, q_1)\| \, \|R_2\| + \gamma_1\gamma_2 \, \|R_1\| \, \|R_2\|\right)$$

Let us take $c_2 < c_1, t_2 < t_1$ such that

$$4d^2(c_2+t_2)\kappa^2\varepsilon^{-4} < \frac{1}{2}\,. \tag{4.28}$$

Because of (ii) in Theorem 2.1, $\|H_{pp}(x_0, q_i)\| \le \gamma_i$. Hence

$$\begin{aligned}\|R\| &\le d(c+t)\big(\|H_{pp}(x_0,q_1)\|\kappa + \|H_{pp}(x_0,q_2)\|\kappa + (c+t)\gamma_1\gamma_2\kappa^2\big)\\ &\le d(c+t)(\gamma_1\kappa + \gamma_2\kappa + 2\gamma_1\gamma_2\kappa^2)\\ &\le 4d(c+t)\gamma_1\gamma_2\kappa^2\\ &= 4d(c+t)\kappa^2\big(\varepsilon^{-1} + |h''(|q_1|)|\,\big)\big(\varepsilon^{-1} + |h''(|q_2|)|\,\big)\,,\end{aligned}$$

where we have used $\kappa, \gamma_1, \gamma_2 > 1$, $c + t < c_1 + t_1 < 2$. Due to this estimate and the fact that, by (i) Theorem 2.1,

$$\begin{aligned}\int_0^1\int_0^1 H_{pp}(x_0,q_1)H_{pp}(x_0,q_2)\,dsd\tau &= \Big(\int_0^1 H_{pp}(x_0,q_1)\,ds\Big)^2\\ &\ge \varepsilon^2\Big(\int_0^1 \big(1+|h''(|q_1|)|\,\big)\,ds\Big)^2 E_d\end{aligned} \tag{4.29}$$

we get

$$\begin{aligned}\int_0^1\int_0^1 \|R\|\,dsd\tau\, E_d &\le 4d(c+t)\kappa^2\varepsilon^{-2}\Big(\int_0^1 \big(1+|h''(|q_1|)|\,\big)\,ds\Big)^2 E_d\\ &\le 4d(c+t)\kappa^2\varepsilon^{-4}\int_0^1\int_0^1 H_{pp}(x_0,q_1)H_{pp}(x_0,q_2)\,dsd\tau\,.\end{aligned} \tag{4.30}$$

Substituting (4.28) in (4.30) and using the elementary formula (recall (4.1))

$$-d\|B\|E_d \le B \le d\|B\|E_d \quad \forall B \in \mathbb{R}^{d\times d}$$

we obtain

$$\begin{aligned}\int_0^1\int_0^1 R\,dsd\tau &\ge -d\int_0^1\int_0^1 \|R\|\,dsd\tau\, E_d\\ &\ge -\frac{1}{2}\int_0^1\int_0^1 H_{pp}(x_0,q_1)H_{pp}(x_0,q_2)\,dsd\tau\,.\end{aligned} \tag{4.31}$$

Then combining (4.27), (4.31) and using again (4.29) we arrive at

$$\begin{aligned}&\int_0^1\int_0^1 \Big(\frac{\partial X(t,x_0,q_1)}{\partial p_0}\Big)^T \frac{\partial X(t,x_0,q_2)}{\partial p_0}\,dsd\tau\\ &\ge \frac{t^2}{2}\int_0^1\int_0^1 H_{pp}(x_0,q_1)H_{pp}(x_0,q_2)\,dsd\tau\\ &\ge \frac{t^2\varepsilon^2}{2}\,E_d\,.\end{aligned}$$

Combining this and (4.26) we complete the proof. □

Proof of Theorem 2.1. Let t_2, c_2 be as in Corollary 4.2, M_4 is from Lemma 3.3, $\varepsilon > 0$ is from the conditions of Theorem 2.1.

Let us take

$$M_7 := \max_{|p|\le M_1} \Big[\big(\varepsilon^{-1} + |h'(|p|)|\big)^2 + \big(\varepsilon^{-1} + |h''(|p|)|\big)\big(\varepsilon^{-1} + |h(|p|)|\big)\Big] , \tag{4.32}$$

$$\rho := \min\big\{\frac{\varepsilon}{8d(4M_4^2 + M_7)} , \ c_2\big\} , \tag{4.33}$$

$$T := \min\big\{\frac{\varepsilon\rho}{4M_4(|h'(0)|^2 + 1)} , \ \frac{\varepsilon\rho}{2M_2} , \rho, t_2\big\} , \tag{4.34}$$

$$r := \frac{\varepsilon\rho}{2} . \tag{4.35}$$

Let $W_t := V_{\rho,t}$ (recall (3.8)) and

$$\mathfrak{D} : p_0 \to X(t, x_0, p_0), \quad \mathfrak{D} : W_t \to \mathbb{R}^d \tag{4.36}$$

Step 1. We will show that $B_r(x_0) \subset \mathfrak{D}(W_t)$. Using (4.6), (4.7), (iii) in Lemma 3.2 and the fact that $\dot X(0) = H_p(x_0, p_0)$, we have

$$\begin{aligned} |X(t) - x_0| &\ge t|\dot X(0)| - \int_0^t (t-\tau)\,|\ddot X(\tau)|\,d\tau \\ &\ge t|H_p(x_0,p_0)| - \int_0^t (t-\tau)\,|A(X(\tau),P(\tau))|\,d\tau \\ &\ge \varepsilon t|h'(|p_0|)| - tM_2 - t^2 \max_{0\le\tau\le t} |A(X(\tau),P(\tau))| . \end{aligned} \tag{4.37}$$

Due to (4.32),

$$\begin{aligned} \|A(x,p)\| &\le d\|H_{xp}\|\,\|H_p\| + d\|H_{pp}\|\,\|H_x\| \\ &\le d\,\big[\,\big(\varepsilon^{-1} + |h'(|p|)|\big)^2 + \big(\varepsilon^{-1} + |h''(|p|)|\big)\big(\varepsilon^{-1} + |h(|p|)|\big)\,\big] \\ &\le dM_7 \end{aligned}$$

for any $x \in \mathbb{R}^d$, $|p| \le M_1$. Piecing together this and (4.15) we obtain

$$\|A(x,p)\| \le 2d\,|h'(|p|)|^2 + dM_7$$

for all $x, p \in \mathbb{R}^n$. This and (3.9) imply

$$\begin{aligned} |A(X(\tau),P(\tau))| &\le 2d\,M_4^2(|h'(|p_0|)| + 1)^2 + dM_7 \\ &\le 4d\,M_4^2(|h'(|p_0|)|^2 + 1) + dM_7 \\ &\le d\,(4M_4^2 + M_7)(|h'(|p_0|)|^2 + 1) . \end{aligned} \tag{4.38}$$

Substituting (4.38) into (4.37) and using $2tM_2 < \varepsilon\rho$, $\ t|h'(|p_0|)| = \rho$ we obtain

$$\begin{aligned} |X(t) - x_0| &\ge \varepsilon t|h'(|p_0|)| - tM_2 - t^2 d(4M_4^2 + M_7)(|h'(|p_0|)|^2 + 1) \\ &\ge \frac{\varepsilon\rho}{2} - d(4M_4^2 + M_7)(\rho^2 + t^2) \end{aligned}$$

for $p_0 \in \partial W_t$, $x_0 \in \mathbb{R}^d$. Since $t < T < \rho$ and $\varepsilon > 8d(4M_4^2 + M_7)\rho$, it follows that

$$|X(t) - x_0| > \frac{\varepsilon\rho}{2} - 2d(4M_4^2 + M_7)\rho^2 > \frac{\varepsilon\rho}{4}$$

or

$$\min_{p_0\in\partial W_t} |X(t, x_0, p_0) - x_0| > r \tag{4.39}$$

for any $x_0 \in \mathbb{R}^d$. On the other hand, using (3.10) with $p_0 = 0$ and (4.34) we have

$$|X(t, x_0, 0) - x_0| \le tM_4(|h'(0)| + 1) < TM_4(|h'(0)| + 1) < \frac{\varepsilon\rho}{4} = r\,. \tag{4.40}$$

Combining (4.39), (4.40) shows that $x_0 \in \mathfrak{D}(W_t)$. Moreover, $B_r(x_0) \subset \mathfrak{D}(W_t)$.

Step 2. Corollary 4.2 implies that $\mathfrak{D}$ is injective. It suffices to notice that, by (4.3) and (i) Theorem 2.1, $\partial X/\partial p_0 \neq 0$ and so the implicit function theorem implies that $\mathfrak{D}$ is a local diffeomorphism for any $t < T$. □

5. Proof of Lemma 2.2

For all $x_0, x \in B_r(x_0)$ and $t < T$ we define the *action function* by

$$\sigma(t, x_0, p_0) = \int_0^t P(\tau)\, dX(\tau) - H(X(\tau), P(\tau))\, d\tau\,. \tag{5.1}$$

Recall that $X(\tau) = X(\tau, x_0, p_0)$, $P(\tau) = P(\tau, x_0, p_0)$ is solution for (1.4) with initial conditions $X(0) = x_0, P(0) = p_0$ and $p_0(t, x, x_0)$ is defined from the equation $X(t, x_0, p_0(t, x, x_0)) = x$. For any fixed $x_0 \in \mathbb{R}^d$ we set

$$p(t, x) := P(t, x_0, p_0(t, x, x_0))\,.$$

The convexity of $H(x, p)$ with respect to the second variable and (1.3) imply

$$L(x, v) = \sup_{p\in\mathbb{R}^d} (pv - H(x, p))\,.$$

Hence

$$I_y(t, x, x_0) = \int_0^t L(y(\tau), \dot y(\tau))\, d\tau \ge \int_0^t [\dot y(\tau) p(\tau, y(\tau)) - H(y(\tau), p(\tau, y(\tau)))]\, d\tau \tag{5.2}$$

for any smooth trajectory $y(\tau)$ with boundary conditions $y(0) = x_0$, $y(t) = x$. However for $y(\tau) = X(\tau)$ we have $p(\tau, y(\tau)) = P(\tau)$. Thus the r.h.s. of (5.2) is equal to $\sigma(t, x_0, p_0(t, x, x_0))$. This and (2.5) yield

$$S(t, x, x_0) = \sigma(t, x_0, p_0(t, x, x_0))\,.$$

Lemma 5.1. *For each $t \in (0, T)$, $x_0 \in \mathbb{R}^d$ we have*

$$S(t, x_0, x_0) \le M_8 \tag{5.3}$$

for some constant $M_8 > 0$.

Proof. Using (4.25) with $p_1 = 0$ and $p_2 = p_0(t, x_0, x_0)$ gives

$$\begin{aligned} |X(t, x_0, 0) - x_0|^2 &= |X(t, x_0, 0) - X(t, x_0, p_0(t, x_0, x_0))|^2 \\ &\ge \frac{t^2\varepsilon^2}{2} |p_0(t, x_0, x_0)|^2\,. \end{aligned} \tag{5.4}$$

Due to (3.10),

$$|X(t, x_0, 0) - x_0|^2 \le t^2 M_4^2 \big(1 + |h'(|0|)|\big)^2\,.$$

Combining this and (5.4) gives

$$|p_0(t, x_0, x_0)|^2 \le \frac{2}{\varepsilon^2} M_4^2 \big(1 + |h'(|0|)|\big)^2. \tag{5.5}$$

On the other hand, an application of (ii) in Theorem 2.1 and (3.23) to (5.1) shows that $\sigma(t, x_0, p_0)$ is bounded if $|p_0|$ is bounded. This remark together with (5.5) completes the proof. □

Proof of Lemma 2.2. Let us choose $M_9 > 0$ such that

$$\frac{M_9}{M_4} - M_5 \geq |h'(M_1)|, \tag{5.6}$$

$$M_{10} := \frac{\varepsilon}{2}\,\omega\big(\frac{M_9}{M_4} - M_5\big)\big(\frac{M_9}{M_4} - M_5\big) - M_6 > 0\,. \tag{5.7}$$

We take $t \in (0,T)$, $y, y_0 \in \mathbb{R}^d$ such that $|y - y_0| = r$. Using

$$\frac{\partial S(t,y,y_0)}{\partial y_0} = -p_0(t,y,y_0)$$

and applying the elementary formula $\phi(1) = \phi(0) + \int_0^1 \phi'(\tau)\,d\tau$ to

$$\phi(\tau) := S(t,y,y+\tau(y_0-y)),$$

$\tau \in [0,1]$ we get

$$\begin{aligned} S(t,y,y_0) &= S(t,y,y) + \int_0^1 (p_0(t,y,y+\tau(y_0-y)), y-y_0)\ d\tau \\ &= S(t,y,y) + \int_0^1 (p_0^\tau, y-\xi_\tau)\ \frac{d\tau}{\tau}\,, \end{aligned} \tag{5.8}$$

where $\xi_\tau = y + \tau(y_0 - y)$ and $p_0^\tau = p_0(t,y,\xi_\tau)$. By (i) and (iii) in Lemma 3.4, we have

$$|p_0^\tau| \geq \omega\big(\frac{r\tau}{tM_4} - M_5\big) \quad \text{provided that } \frac{r\tau}{tM_4} - M_5 > M_4\,, \tag{5.9}$$

$$(p_0^\tau, y-\xi_\tau) \geq \frac{\varepsilon}{2}\,|p_0^\tau|\big(\frac{r\tau}{M_4} - tM_5\big) - tM_6\,, \tag{5.10}$$

where we have used $|y - \xi_\tau| = \tau r$. Setting $s = tM_9/r$ and using (5.9)-(5.10) we get

$$(p_0^\tau, y-\xi_\tau) \geq \frac{\varepsilon}{2}\,|p_0^\tau|\big(t\frac{M_9}{M_4} - tM_5\big) - tM_6 > tM_{10} > 0 \quad \text{for } \tau \in [s,1]\,,$$

where we have used $\tau r \geq tM_9$. Notice that due to (5.6) the condition in (5.9) is satisfied.

If we assume that $t \leq r/(2M_9)$, then $s \leq 1/2$ and so

$$\int_0^1 (p_0^\tau, y-\xi_\tau)\ \frac{d\tau}{\tau} \geq \int_{1/2}^1 - \Big|\int_0^s\Big|\,. \tag{5.11}$$

Using again (5.10) and, due to (i) Lemma 3.4,

$$|p_0^\tau| \geq \omega\big(\frac{\tau r}{tM_4} - M_5\big)$$

we obtain

$$(p_0^\tau, y-\xi_\tau) \geq \frac{\varepsilon}{2}|p_0^\tau|\big(\frac{r}{2M_4} - tM_5\big) - tM_6 \tag{5.12}$$

$$\geq \frac{\varepsilon}{2}\omega\big(\frac{r}{2tM_4} - M_5\big)\big(\frac{r}{2M_4} - tM_5\big) - tM_6 \quad \text{for } \tau \in [\frac{1}{2},1]\,. \tag{5.13}$$

On the other hand, due to (ii) Lemma 3.4 for $0 \leq \tau \leq s$, we have

$$|p_0^\tau| \leq \omega\big(\frac{2}{\varepsilon}\,\frac{r\tau}{t} + M_5\big) \leq \omega\big(\frac{2}{\varepsilon}\,M_9 + M_5\big) =: M_{11}$$

and so

$$\Big|\int_0^s (p_0^\tau, y-\xi_\tau)\ \frac{d\tau}{\tau}\Big| \leq r\int_0^s |p_0^\tau|\,d\tau \leq rM_{11}\,. \tag{5.14}$$

Combining (5.8), (5.11), (5.12), (5.14) we arrive at

$$\begin{aligned}&\min_{|y-y_0|=r} S(t,y,y_0)\\ &\geq -M_8 + \Big[\frac{\varepsilon}{2}\omega\big(\frac{r}{2tM_4} - M_5\big)\big(\frac{r}{2M_4} - tM_5\big) - tM_6\Big]\ln 2 - rM_{11}\end{aligned} \tag{5.15}$$

for $t < r/(2M_9)$.

Step 2. We choose $r_1 < r$ such that

$$2r_1 < \frac{\varepsilon r \ln 2}{16M_4}, \quad \frac{4r_1}{\varepsilon} < \frac{r}{4M_4}. \tag{5.16}$$

Using (5.8) and (i) Lemma 3.4, we get

$$\begin{aligned}\max_{|y-y_0|=r_1} S(t,y,y_0) &\leq M_8 + r_1\int_0^1 |p_0(t,y,y+\tau(y-y_0))|\,d\tau\\ &\leq M_8 + r_1\omega\big(\frac{2r_1}{\varepsilon t} + M_5\big)\end{aligned} \tag{5.17}$$

for $y, y_0 \in \mathbb{R}^d$ such that $|y-y_0| = r_1$.

Since $\omega(z)$ is increasing function and $\omega(z) \to +\infty$, then there exists $T_1 < r/(2M_9)$ such that for all $t < T_1$

$$\text{right-hand side of (5.15)} \geq \frac{\varepsilon r\ln 2}{16M_4}\omega\big(\frac{r}{4tM_4}\big) \tag{5.18}$$

and

$$\text{right-hand side of (5.17)} \leq 2r_1\omega\big(\frac{4r_1}{\varepsilon t}\big). \tag{5.19}$$

Applying (5.16), (5.18), (5.19) to (5.15) and (5.17) we arrive at (2.6). Finally we notice that (2.7) is a direct consequence of (5.15). □

References

[1] Akhiezer, N. I., *The calculus of variations.* Blaisdell publishing company, New York, London, 1962.

[2] Ball, J. M., The calculus of variations and materials science. *Q. Appl. Math.* **56**, 719-740 (1998).

[3] Giaquinta, M., Hildebrandt, S., *Calculus of variations.* (Grundlehren Math. Wiss., **310** and **311**), Springer, Berlin, 1996.

[4] Kolokoltsov, V. N., *Semiclassical Analysis for Diffusions and Stochastic Processes.* (Lecture Notes in Mathematics, **1724**), Springer, Berlin, 2000.

[5] Kolokoltsov, V. N., Stochastic Hamilton-Jacobi-Bellman equation and stochastic Hamiltonian systems. *J. Dyn. and Control Syst.* **2**, 299-379 (1996).

[6] Kolokoltsov, V. N., Schilling, R. L., Tyukov, A. E., Estimates for multiple stochastic integrals and stochastic Hamilton-Jacobi equations. *Rev. Mat. Iberoamericana* **20**(2), 333-380 (2004).

[7] Maslov, V. P., *Méthodes opératorielles.* Mir, Moscow, 1987 (In French).

[8] Tihomirov, V. M., Convex analysis. In: Gamkrelidze (ed.), *Analysis 2. Encyclopedia of mathematical sciences,* **14**, pp. 1-92, Springer, 1990.

Vassili N. Kolokol'tsov
Dept of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail address: v.kolokoltsov@warwick.ac.uk

Alexey E. Tyukov
Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, United Kingdom
E-mail address: tyukovA@cardiff.ac.uk, tyukov@yahoo.co.uk