Ergod. Th. & Dynam. Sys. (2008), **28**, 267–290
doi: 10.1017/S0143385707000351

On the total disconnectedness of the quotient Aubry set

ALFONSO SORRENTINO

Department of Mathematics, Princeton University, Washington Road, Princeton, NJ 08544, USA
(e-mail: asorrent@math.princeton.edu)

(*Received* 19 *July* 2006 *and accepted in revised form* 2 *April* 2007)

Abstract. In this paper we show that the quotient Aubry set, associated to a sufficiently smooth mechanical or symmetrical Lagrangian, is totally disconnected (i.e. every connected component consists of a single point). This result is optimal, in the sense of the regularity of the Lagrangian, as Mather's counterexamples (J. N. Mather. Examples of Aubry sets. *Ergod. Th. & Dynam. Sys.* **24**(5) (2004), 1667–1723) show. Moreover, we discuss the relation between this problem and a Morse–Sard-type property for (the difference of) critical subsolutions of Hamilton–Jacobi equations.

1. *Introduction*

In Mather's studies of the dynamics of Lagrangian systems and the existence of Arnold diffusion, it turns out that understanding certain aspects of the *Aubry set* and, in particular, what is called the *quotient Aubry set*, may help in the construction of orbits with interesting behavior.

While in the case of twist maps (see, for instance, [**2, 11**] and references therein) there is a detailed structure theory for these sets, in more degrees of freedom only a few results are known. In particular, it seems to be useful to know whether the quotient Aubry set is 'small' in some sense of dimension (e.g., vanishing topological or box dimension).

In [**16**] Mather showed that if the state space has dimension ≤ 2 (in the non-autonomous case) or the Lagrangian is the kinetic energy associated to a Riemannian metric and the state space has dimension ≤ 3, then the quotient Aubry set is totally disconnected, i.e. every connected component consists of a single point (in a compact metric space this is equivalent to vanishing topological dimension). In the autonomous case, with $\dim M \leq 3$, the same argument shows that this quotient is totally disconnected as long as the Aubry set does not intersect the zero section of $\mathrm{T}M$ (this is the case when the cohomology class is large enough in norm).

What happens in higher dimensions? Unfortunately, this is generally not true. In fact, Burago *et al* in [**5**] provided an example that does not satisfy this property (they did not discuss it in their work, but it follows from the results therein). More strikingly, Mather provided in [**17**] several examples of quotient Aubry sets that are not only non-totally-disconnected, but even isometric to closed intervals. All these examples come from mechanical Lagrangians on $\mathrm{T}\mathbb{T}^d$ (i.e. the sum of the kinetic energy and a potential) with $d \geq 3$. In particular, for every $\varepsilon > 0$, he provided a potential $U \in C^{2d-3,1-\varepsilon}(\mathbb{T}^d)$, whose associated quotient Aubry set is isometric to an interval. As the author himself noticed, it is not possible to improve the differentiability of these examples, due to the construction carried out.

The main aim of this article is to show that the counterexamples provided by Mather are optimal, in the sense that for more regular mechanical Lagrangians the associated quotient Aubry sets—corresponding to the zero cohomology class—are totally disconnected.

In particular, our result will also apply to slightly more general Lagrangians, satisfying certain conditions on the zero section; in this case, we shall be able to show that the quotient Aubry set, corresponding to a well-specified cohomology class, is totally disconnected.

We shall also outline a possible approach to generalize this result, pointing out how it is related to a *Morse–Sard-type* problem; from this and Sard's lemma, one can easily recover Mather's result in dimension $d = 2$ (autonomous case).

It is important to point out that most of this approach has been inspired by Albert Fathi's talk [**8**], in which he used this relation with Sard's lemma to show a simpler way to construct mechanical Lagrangians on $\mathrm{T}\mathbb{T}^N$, whose quotient Aubry sets are Lipschitz equivalent to any given *doubling* metric space or, equivalently, to any space with finite *Assouad dimension* (see [**12**] for a similar construction). In this case we do not get a neat relation between their regularity and N, as in Mather's work, but we can only observe that N goes to infinity as r increases. It would be interesting to study in depth the relation between the dimension of the quotient Aubry set, the regularity of the Lagrangian and the dimension of the state space. Our result may be seen as a first step in this direction.

Post Scriptum. Just before submitting this paper, we learnt that analogous results had been proven independently by Albert Fathi, Alessio Figalli and Ludovic Rifford, using a similar approach (to be published).

Moreover, in '*A generic property of families of Lagrangian systems*' (to appear in *Annals of Mathematics*), Patrick Bernard and Gonzalo Contreras managed to show that generically, in Mañé's sense, there are at most $1 + \dim H^1(M; \mathbb{R})$ ergodic minimizing measures, for each cohomology class $c \in H^1(M; \mathbb{R})$. As a corollary of this striking result, one gets that generically the quotient Aubry set is finite for each cohomology class and it consists of at most $1 + \dim H^1(M; \mathbb{R})$ elements.

2. *The Aubry set and the quotient Aubry set*

Let M be a compact and connected smooth manifold without boundary. Denote by $\mathrm{T}M$ its tangent bundle and T^*M the cotangent one. A point of $\mathrm{T}M$ will be denoted by (x, v), where $x \in M$ and $v \in \mathrm{T}_x M$, and a point of T^*M by (x, p), where $p \in \mathrm{T}_x^* M$ is a linear form on the vector space $\mathrm{T}_x M$. Let us fix a Riemannian metric g on it and denote with d the induced metric on M; let $\|\cdot\|_x$ be the norm induced by g on $\mathrm{T}_x M$; we shall use the same notation for the norm induced on $\mathrm{T}_x^* M$.

Definition. A function $L : \mathrm{T}M \longrightarrow \mathbb{R}$ is called a *Tonelli Lagrangian* if:

(i) $L \in C^2(\mathrm{T}M)$;

(ii) L is strictly convex in the fibers, i.e. the second partial vertical derivative $\partial^2 L(x, v)/\partial v^2$ is positive definite, as a quadratic form, for any $(x, v) \in \mathrm{T}M$;

(iii) L is superlinear in each fiber, i.e.

$$\lim_{\|v\|_x \to +\infty} \frac{L(x, v)}{\|v\|_x} = +\infty$$

(this condition is independent of the choice of the Riemannian metric).

Given a Lagrangian, we can define the associated *Hamiltonian* as a function on the cotangent bundle:

$$\begin{aligned} H : \mathrm{T}^*M &\longrightarrow \mathbb{R} \\ (x, p) &\longmapsto \sup_{v \in \mathrm{T}_x M} \{\langle p, v\rangle_x - L(x, v)\} \end{aligned}$$

where $\langle \cdot, \cdot \rangle_x$ represents the canonical pairing between the tangent and cotangent space.

If L is a Tonelli Lagrangian, one can easily prove that H is finite everywhere, C^2, superlinear and strictly convex in the fibers. Moreover, under the above assumptions, one can define a diffeomorphism between $\mathrm{T}M$ and T^*M, called the *Legendre transform*:

$$\begin{aligned} \mathcal{L} : \mathrm{T}M &\longrightarrow \mathrm{T}^*M \\ (x, v) &\longmapsto \left(x, \frac{\partial L}{\partial v}(x, v)\right). \end{aligned}$$

In particular, $\mathcal{L}$ is a conjugation between the two flows (namely the Euler–Lagrange and Hamiltonian flows) and

$$H \circ \mathcal{L}(x, v) = \left\langle \frac{\partial L}{\partial v}(x, v), v \right\rangle_x - L(x, v).$$

Observe that if η is a 1-form on M, then we can define a function on the tangent space

$$\begin{aligned} \hat{\eta} : \mathrm{T}M &\longrightarrow \mathbb{R} \\ (x, v) &\longmapsto \langle \eta(x), v\rangle_x \end{aligned}$$

and consider a new Tonelli Lagrangian $L_\eta = L - \hat{\eta}$. The associated Hamiltonian will be $H_\eta(x, p) = H(x, p + \eta)$. Moreover, if η is closed, then $\int L\, dt$ and $\int L_\eta\, dt$ have the same extremals and therefore the Euler–Lagrange flows on $\mathrm{T}M$ associated to L and L_η are the same.

Although the extremals are the same, this is not generally true for the minimizers. What one can say is that they stay the same when we change the Lagrangian by an exact 1-form. Thus, for fixed L, the minimizers depend only on the de Rham cohomology class $c = [\eta] \in H^1(M; \mathbb{R})$. From here comes the interest in considering modified Lagrangians, corresponding to different cohomology classes.

Let us fix η, a smooth (C^2 is enough for what follows) 1-form on M, and let $c = [\eta] \in H^1(M; \mathbb{R})$ be its cohomology class.

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

As done by Mather in [**15**], it is convenient to introduce, for $t > 0$ and $x, y \in M$, the following quantity:

$$h_{\eta,t}(x, y) = \inf \int_0^t L_\eta(\gamma(s), \dot{\gamma}(s))\, ds,$$

where the infimum is taken over all piecewise C^1 paths $\gamma : [0, t] \longrightarrow M$, such that $\gamma(0) = x$ and $\gamma(t) = y$. We define the *Peierls barrier* as

$$h_\eta(x, y) = \liminf_{t\to+\infty} (h_{\eta,t}(x, y) + \alpha(c)t),$$

where $\alpha : H^1(M; \mathbb{R}) \longrightarrow \mathbb{R}$ is Mather's α function (see [**14**]). It can be shown that this function is convex and that (only for the autonomous case) the lim inf can be replaced by lim.

Observe that h_η does not depend only on the cohomology class c, but also on the choice of the representant; namely, if $\eta' = \eta + df$, then $h_{\eta'}(x, y) = h_\eta(x, y) + f(y) - f(x)$.

PROPOSITION 1. *The values of the map h_η are finite. Moreover, the following properties hold:*
(i) *h_η is Lipschitz;*
(ii) *for each $x \in M$, $h_\eta(x, x) \geq 0$;*
(iii) *for each $x, y, z \in M$, $h_\eta(x, y) \leq h_\eta(x, z) + h_\eta(z, y)$;*
(iv) *for each $x, y \in M$, $h_\eta(x, y) + h_\eta(y, x) \geq 0$.*

For a proof of the above claims and more, see [**7, 9, 15**]. Inspired by these properties, we can define

$$\begin{aligned} \delta_c :\ M \times M &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto h_\eta(x, y) + h_\eta(y, x) \end{aligned}$$

(observe that this function does actually depend only on the cohomology class).

This function is positive, symmetric and satisfies the triangle inequality; therefore, it is a pseudometric on

$$\mathcal{A}_{L,c} = \{x \in M : \delta_c(x, x) = 0\}.$$

$\mathcal{A}_{L,c}$ is called the *Aubry set* (or *projected Aubry set*) associated to L and c, and δ_c is *Mather's pseudometric*. In [**15**], Mather has showed that this is a non-empty compact subset of M that can be Lipschitzly lifted to a compact invariant subset of TM.

Definition. The *quotient Aubry set* $(\bar{\mathcal{A}}_{L,c}, \bar{\delta}_c)$ is the metric space obtained by identifying two points in $\mathcal{A}_{L,c}$, if their δ_c-pseudodistance is zero.

We shall denote an element of this quotient by $\bar{x} = \{y \in \mathcal{A}_{L,c} :\ \delta_c(x, y) = 0\}$. These elements (that are also called *c-static classes*, see [**7**]) provide a partition of $\mathcal{A}_{L,c}$ into compact subsets that can be lifted to invariant subsets of TM. They are really interesting from a dynamical systems point of view, since they contain the α and ω limit sets of c-minimizing orbits (see [**7, 15**] for more details).

For the sake of our proof, it is convenient to adopt Fathi's *weak KAM theory* point of view (we point the reader to [**9**] for a self-contained presentation).

Definition. A locally Lipschitz function $u : M \longrightarrow \mathbb{R}$ is a *subsolution* of $H_\eta(x, d_x u) = k$, with $k \in \mathbb{R}$, if $H_\eta(x, d_x u) \le k$ for almost every $x \in M$.

This definition makes sense, because, by Rademacher's theorem, we know that $d_x u$ exists almost everywhere.

It is possible to show that there exists $c[H_\eta] \in \mathbb{R}$, such that $H_\eta(x, d_x u) = k$ admits no subsolutions for $k < c[H_\eta]$ and has subsolutions for $k \ge c[H_\eta]$. The constant $c[H_\eta]$ is called *Mañé's critical value* and coincides with $\alpha(c)$, where $c = [\eta]$ (see [**7**]).

Definition. $u : M \longrightarrow \mathbb{R}$ is a *η-critical subsolution*, if $H_\eta(x, d_x u) \le \alpha(c)$ for almost every $x \in M$.

Denote by $\mathcal{S}_\eta$ the set of critical subsolutions. This set $\mathcal{S}_\eta$ is non-empty. In fact, Fathi showed (see [**9**]) the following proposition.

PROPOSITION 2. *If $u : M \longrightarrow \mathbb{R}$ is a η-critical subsolution, then for every $x, y \in M$,*

$$u(y) - u(x) \le h_\eta(x, y).$$

Moreover, for any $x \in M$, the function $h_{\eta,x}(\cdot) := h_\eta(x, \cdot)$ is a η-critical subsolution.

Using this result, he provided a nice representation of h_η, in terms of the η-critical subsolutions.

COROLLARY 1. *If $x \in \mathcal{A}_{L,c}$ and $y \in M$,*

$$h_\eta(x, y) = \sup_{u \in \mathcal{S}_\eta} (u(y) - u(x)).$$

This supremum is actually attained.

Proof. It is clear, from the proposition above, that

$$h_\eta(x, y) \ge \sup_{u \in \mathcal{S}_\eta} (u(y) - u(x)).$$

Let us show the other inequality. In fact, since $h_{\eta,x}$ is a η-critical subsolution and $x \in \mathcal{A}_{L,c}$ (i.e. $h_\eta(x, x) = 0$), then

$$h_\eta(x, y) = h_{\eta,x}(y) - h_{\eta,x}(x) \le \sup_{u \in \mathcal{S}_\eta} (u(y) - u(x)).$$

This shows that the supremum is attained. □

This result can be still improved. Fathi and Siconolfi proved the following in [**10**].

THEOREM. (Fathi and Siconolfi) *For any η-critical subsolution $u : M \longrightarrow \mathbb{R}$ and for each $\varepsilon > 0$, there exists a C^1 function $\tilde{u} : M \longrightarrow \mathbb{R}$ such that:*

(i) $\tilde{u}(x) = u(x)$ *and* $H_\eta(x, d_x \tilde{u}) = \alpha(c)$ *on* $\mathcal{A}_{L,c}$;

(ii) $|\tilde{u}(x) - u(x)| < \varepsilon$ *and* $H_\eta(x, d_x \tilde{u}) < \alpha(c)$ *on* $M \setminus \mathcal{A}_{L,c}$.

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

In particular, this implies that C^1 η-critical subsolutions are dense in $\mathcal{S}_\eta$ with the uniform topology. This result has been recently improved by Bernard (see [**4**]), showing that every η-critical subsolution coincides, on the Aubry set, with a $C^{1,1}$ η-critical subsolution.

Denote the set of C^1 η-critical subsolutions by $\mathcal{S}_\eta^1$ and the set of $C^{1,1}$ η-critical subsolutions by $\mathcal{S}_\eta^{1,1}$.

COROLLARY 2. *For $x, y \in \mathcal{A}_{L,c}$, the following representation holds:*

$$h_\eta(x, y) = \sup_{u \in \mathcal{S}_\eta^1} (u(y) - u(x)) = \sup_{u \in \mathcal{S}_\eta^{1,1}} (u(y) - u(x)).$$

Moreover, these suprema are attained.

It turns out to be convenient to characterize the elements of $\bar{\mathcal{A}}_{L,c}$ (i.e. the c-quotient classes) in terms of η-critical subsolutions.

Let us consider the following set:

$$\mathcal{D}_c = \{u - v : u, v \in \mathcal{S}_\eta\}$$

(it depends only on the cohomology class c and not on η) and denote by $\mathcal{D}_c^1$ and $\mathcal{D}_c^{1,1}$ the sets corresponding, respectively, to C^1 and $C^{1,1}$ η-critical subsolutions.

PROPOSITION 3. *For $x, y \in \mathcal{A}_{L,c}$,*

$$\begin{aligned}\delta_c(x, y) &= \sup_{w \in \mathcal{D}_c} (w(y) - w(x)) = \sup_{w \in \mathcal{D}_c^1} (w(y) - w(x)) \\ &= \sup_{w \in \mathcal{D}_c^{1,1}} (w(y) - w(x))\end{aligned}$$

and these suprema are attained.

Proof. From the definition of $\delta_c(x, y)$, we immediately get

$$\begin{aligned}\delta_c(x, y) &= h_\eta(x, y) + h_\eta(y, x) \\ &= \sup_{u \in \mathcal{S}_\eta} (u(y) - u(x)) + \sup_{v \in \mathcal{S}_\eta} (v(x) - v(y)) \\ &= \sup_{u, v \in \mathcal{S}_\eta} [(u(y) - v(y)) - (u(x) - v(x))] \\ &= \sup_{w \in \mathcal{D}_c} (w(y) - w(x)).\end{aligned}$$

The other equalities follow from the density results we mentioned above. □

PROPOSITION 4. *If $w \in \mathcal{D}_c$, then $d_x w = 0$ on $\mathcal{A}_{L,c}$. Therefore, $\mathcal{A}_{L,c} \subseteq \bigcap_{w \in \mathcal{D}_c^{1,1}} \mathrm{Crit}(w)$, where $\mathrm{Crit}(w)$ is the set of critical points of w.*

Proof. This is an immediate consequence of a result by Fathi (see [**9**]); namely, if $u, v \in \mathcal{S}_\eta$, then they are differentiable on $\mathcal{A}_{L,c}$ and $d_x u = d_x v$. □

PROPOSITION 5. *If $w \in \mathcal{D}_c$, then it is constant on any quotient class of $\bar{\mathcal{A}}_{L,c}$; namely, if $x, y \in \mathcal{A}_{L,c}$ and $\delta_c(x, y) = 0$, then $w(x) = w(y)$.*

Proof. From the representation formula above, it follows that

$$0 = \delta_c(x, y) = \sup_{\tilde{w} \in \mathcal{D}_c} (\tilde{w}(y) - \tilde{w}(x)) \geq w(y) - w(x),$$
$$0 = \delta_c(y, x) = \sup_{\tilde{w} \in \mathcal{D}_c} (\tilde{w}(x) - \tilde{w}(y)) \geq w(x) - w(y). \qquad \square$$

For any $w \in \mathcal{D}_c^1$, let us define the following *evaluation function*:

$$\begin{aligned} \varphi_w : (\bar{\mathcal{A}}_{L,c}, \bar{\delta}_c) &\longrightarrow (\mathbb{R}, |\cdot|) \\ \bar{x} &\longmapsto w(x). \end{aligned}$$

- φ_w is well defined, i.e. it does not depend on the element of the class at which w is evaluated;
- $\varphi_w(\bar{\mathcal{A}}_{L,c}) = w(\mathcal{A}_{L,c}) \subseteq w(\mathrm{Crit}(w))$;
- φ_w is Lipschitz, with Lipschitz constant 1. In fact,

$$\varphi_w(\bar{x}) - \varphi_w(\bar{y}) = w(x) - w(y) \leq \delta_c(x, y) = \bar{\delta}_c(\bar{x}, \bar{y}),$$
$$\varphi_w(\bar{y}) - \varphi_w(\bar{x}) = w(y) - w(x) \leq \delta_c(y, x) = \bar{\delta}_c(\bar{y}, \bar{x}).$$

Therefore,

$$|\varphi_w(\bar{x}) - \varphi_w(\bar{y})| \leq \bar{\delta}_c(\bar{x}, \bar{y}).$$

As we shall see, these functions play a key role in the proof of our result.

3. *The main result*

Our main goal is to show that, under suitable hypotheses on L, there is a well-specified cohomology class c_L, for which $(\bar{\mathcal{A}}_{L,c_L}, \bar{\delta}_{c_L})$ is totally disconnected, i.e. every connected component consists of a single point.

Consider $L : TM \longrightarrow \mathbb{R}$ a Tonelli Lagrangian and the associated Legendre transform

$$\begin{aligned} \mathcal{L} : \mathrm{T}M &\longrightarrow \mathrm{T}^*M \\ (x, v) &\longmapsto \left(x, \frac{\partial L}{\partial v}(x, v)\right). \end{aligned}$$

Remember that T^*M, as a cotangent bundle, may be equipped with a *canonical* symplectic structure. Namely, if $(\mathcal{U}, x_1, \ldots, x_d)$ is a local coordinate chart for M and $(\mathrm{T}^*\mathcal{U}, x_1, \ldots, x_d, p_1, \ldots, p_d)$ are the associated cotangent coordinates, one can define the 2-form

$$\omega = \sum_{i=1}^{d} dx_i \wedge dp_i.$$

It is easy to show that ω is a symplectic form (i.e. it is non-degenerate and closed). In particular, one can check that ω is intrinsically defined, by considering the 1-form on $\mathrm{T}^*\mathcal{U}$,

$$\lambda = \sum_{i=1}^{d} p_i \, dx_i,$$

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

which satisfies $\omega = -d\lambda$ and is coordinate independent; in fact, in terms of the natural projection

$$\pi : \mathrm{T}^*M \longrightarrow M$$
$$(x, p) \longmapsto x$$

the form λ may be equivalently defined pointwise without coordinates by

$$\lambda_{(x,p)} = (d\pi_{(x,p)})^* p \in \mathrm{T}^*_{(x,p)}\mathrm{T}^*M.$$

The 1-form λ is called the *Liouville form* (or the *tautological form*).

Consider now the section of T^*M given by

$$\Lambda_L = \mathcal{L}(M \times \{0\}) = \left\{ \left(x, \frac{\partial L}{\partial v}(x, 0) \right) : x \in M \right\},$$

corresponding to the 1-form

$$\eta_L(x) = \frac{\partial L}{\partial v}(x, 0) \cdot dx = \sum_{i=1}^{d} \frac{\partial L}{\partial v_i}(x, 0)\, dx_i.$$

We would like this 1-form to be closed, which is equivalent to asking that Λ_L is a Lagrangian submanifold, in order to consider its cohomology class $c_L = [\eta_L] \in H^1(M; \mathbb{R})$. Observe that this cohomology class can be defined in a more intrinsic way; in fact, consider the projection

$$\pi_{|\Lambda_L} : \Lambda_L \subset \mathrm{T}^*M \longrightarrow M;$$

this induces an isomorphism between the cohomology groups $H^1(M; \mathbb{R})$ and $H^1(\Lambda_L; \mathbb{R})$. The preimage of $[\lambda_{|\Lambda_L}]$ under this isomorphism is called the *Liouville class* of Λ_L and one can easily show that it coincides with c_L.

We can define the set

$$\mathbb{L}(M) = \{L : \mathrm{T}M \longrightarrow \mathbb{R} : \ L \text{ is a Tonelli Lagrangian and } \Lambda_L \text{ is Lagrangian}\}.$$

This set is non-empty and consists of Lagrangians of the form

$$L(x, v) = f(x) + \langle \eta(x), v \rangle_x + O(\|v\|^2),$$

with $f \in C^2(M)$ and η a C^2 closed 1-form on M. In particular, it includes the mechanical Lagrangians, i.e. Lagrangians of the form

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

namely, the sum of the kinetic energy and a potential $U : M \longrightarrow \mathbb{R}$. More generally, it contains the *symmetrical* (or *reversible*) *Lagrangians*, i.e. Lagrangians $L : \mathrm{T}M \longrightarrow \mathbb{R}$ such that

$$L(x, v) = L(x, -v),$$

for every $(x, v) \in \mathrm{T}M$.

In fact, in the above cases, $\partial L(x, 0)/\partial v \equiv 0$; therefore, $\Lambda_L = M \times \{0\}$ (the zero section of the cotangent space), which is clearly Lagrangian, and $c_L = 0$.

We can now state our main result.

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

MAIN THEOREM. *Let M be a compact connected manifold of dimension $d \geq 1$ and let $L \in \mathbb{L}(M)$ be a Lagrangian such that $L(x, 0) \in C^r(M)$, with $r \geq 2d - 2$ and $\partial L(x, 0)/\partial v \in C^2(M)$. Then, the quotient Aubry set $(\bar{\mathcal{A}}_{L,c_L}, \bar{\delta}_{c_L})$, corresponding to the Liouville class of Λ_L, is totally disconnected, i.e. every connected component consists of a single point.*

This result immediately implies the following corollary.

COROLLARY 3. (Symmetrical Lagrangians) *Let M be a compact connected manifold of dimension $d \geq 1$ and let $L(x, v)$ be a symmetrical Tonelli Lagrangian on $\mathrm{T}M$, such that $L(x, 0) \in C^r(M)$, with $r \geq 2d - 2$. Then, the quotient Aubry set $(\bar{\mathcal{A}}_{L,0}, \bar{\delta}_0)$ is totally disconnected.*

More specifically, we have the following.

COROLLARY 4. (Mechanical Lagrangians) *Let M be a compact connected manifold of dimension $d \geq 1$ and let $L(x, v) = \frac{1}{2}\|v\|_x^2 + U(x)$ be a mechanical Lagrangian on $\mathrm{T}M$, such that the potential $U \in C^r(M)$, with $r \geq 2d - 2$. Then, the quotient Aubry set $(\bar{\mathcal{A}}_{L,0}, \bar{\delta}_0)$ is totally disconnected.*

Remark. This result is optimal, in the sense of the regularity asked of the potential U, for $\bar{\mathcal{A}}_{L,0}$ to be totally disconnected. In fact, Mather provided in [**17**] examples of quotient Aubry sets isometric to the unit interval, corresponding to mechanical Lagrangians $L \in C^{2d-3,1-\varepsilon}(\mathrm{T}\mathbb{T}^d)$, for any $0 < \varepsilon < 1$.

Before proving the main theorem, it will be beneficial to show some useful results.

LEMMA 1. *Let us consider $L \in \mathbb{L}(M)$, such that $\partial L(x, 0)/\partial v \in C^2(M)$, and let H be the associated Hamiltonian.*

(1) *Every constant function $u \equiv$ constant is a η_L-critical subsolution. In particular, all η_L-critical subsolutions are such that $d_x u \equiv 0$ on $\mathcal{A}_{L,c_L}$.*

(2) *For every $x \in M$,*

$$\frac{\partial H_{\eta_L}}{\partial p}(x, 0) = \frac{\partial H}{\partial p}(x, \eta_L(x)) = 0.$$

Proof. (1) The second part follows immediately from the fact that, if $u, v \in \mathcal{S}_{\eta_L}$, then they are differentiable on $\mathcal{A}_{L,c_L}$ and $d_x u = d_x v$ (see [**9**]).

Let us show that $u \equiv$ constant is a η_L-critical subsolution; namely, that

$$H_{\eta_L}(x, 0) \leq \alpha(c_L)$$

for every $x \in M$. It is sufficient to observe the following.

- $H_{\eta_L}(x, 0) = -L(x, 0)$; in fact,

$$\begin{aligned} H_{\eta_L}(x, 0) &= H(x, \eta_L(x)) = H\left(x, \frac{\partial L}{\partial v}(x, 0)\right) \\ &= \left\langle \frac{\partial L}{\partial v}(x, 0),\ 0 \right\rangle_x - L(x, 0) \\ &= -L(x, 0). \end{aligned}$$

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

- Let v be *dominated* by $L_{\eta_L} + \alpha(c_L)$ (see [**9**], for the existence of such functions), i.e. for each continuous piecewise C^1 curve $\gamma : [a, b] \longrightarrow M$ we have

$$v(\gamma(b)) - v(\gamma(a)) \leq \int_a^b L_{\eta_L}(\gamma(t), \dot{\gamma}(t))\, dt + \alpha(c_L)(b-a).$$

Then, considering the constant path $\gamma(t) \equiv x$, one can easily deduce that

$$\alpha(c_L) \geq \sup_{x\in M}(-L_{\eta_L}(x, 0)) = -\inf_{x\in M} L_{\eta_L}(x, 0);$$

therefore,

$$\alpha(c_L) \geq -L_{\eta_L}(x, 0) = -L(x, 0) = H_{\eta_L}(x, 0)$$

for every $x \in M$.

(2) The inverse of the Legendre transform can be written in coordinates,

$$\begin{aligned} \mathcal{L}^{-1} : \mathrm{T}^*M &\longrightarrow \mathrm{T}M \\ (x, p) &\longmapsto \left(x, \frac{\partial H}{\partial p}(x, p)\right). \end{aligned}$$

Therefore,

$$\begin{aligned} (x, 0) &= \mathcal{L}^{-1}\Big(\mathcal{L}(x, 0)\Big) = \mathcal{L}^{-1}\left(x, \frac{\partial L}{\partial v}(x, 0)\right) \\ &= \mathcal{L}^{-1}((x, \eta_L(x))) = \left(x, \frac{\partial H}{\partial p}(x, \eta_L(x))\right). \end{aligned}$$ □

In particular, observing that for any η_L-critical subsolution u, $H_{\eta_L}(x, d_x u) = \alpha(c_L)$ on $\mathcal{A}_{L,c_L}$, we can easily deduce from the above that

$$\mathcal{A}_{L,c_L} \subseteq \{L(x, 0) = -\alpha(c_L)\} = \{H(x, \eta_L(x)) = \alpha(c_L)\}$$

and

$$\alpha(c_L) = \sup_{x\in M}(-L(x, 0)) = -\inf_{x\in M} L(x, 0) =: e_0,$$

as denoted in [**6, 13**].

Let us observe that in general

$$e_0 \leq \min_{c\in H^1(M;\mathbb{R})} \alpha(c) = -\beta(0),$$

where $\beta : H_1(M; \mathbb{R}) \longrightarrow \mathbb{R}$ is Mather's β-function, i.e. the convex conjugate of α (in [**6, 13**], the right-hand-side quantity is referred to as *strict critical value*). Therefore, we are considering an extremal case in which $e_0 = \alpha(c_L) = \min \alpha(c)$; it follows also quite easily that $c_L \in \partial\beta(0)$, namely, it is a subgradient of β at zero.

A crucial step in the proof of our result will be the following lemma, which can be read as a sort of relaxed version of Sard's lemma (the proof will be mainly based on the one in [**1**]).

MAIN LEMMA. *Let $U \in C^r(M)$, with $r \geq 2d - 2$, be a non-negative function, vanishing somewhere, and denote $\mathcal{A} = \{U(x) = 0\}$. If $u : M \longrightarrow \mathbb{R}$ is C^1 and satisfies $\|d_x u\|_x^2 \leq U(x)$ in an open neighborhood of $\mathcal{A}$, then $|u(\mathcal{A})| = 0$ (where $|\cdot|$ denotes the Lebesgue measure in $\mathbb{R}$).*

See §4 for its proof.

In particular, it implies this essential property.

COROLLARY 5. *Under the hypotheses of the main theorem, if $u \in \mathcal{S}_{\eta_L}$, then*

$$|u(\mathcal{A}_{L,c_L})| = 0$$

(*where* $|\cdot|$ *denotes the Lebesgue measure in* $\mathbb{R}$).

Proof of Corollary 5. First of all, we can assume that $u \in \mathcal{S}^1_{\eta_L}$, because of Fathi and Siconolfi's theorem. By Taylor's formula, it follows that there exists an open neighborhood W of $\mathcal{A}_{L,c_L}$, such that for all $x \in W$,

$$\alpha(c_L) \geq H_{\eta_L}(x, d_x u) = H_{\eta_L}(x, 0) + \frac{\partial H_{\eta_L}}{\partial p}(x, 0) \cdot d_x u + \int_0^1 (1-t) \frac{\partial^2 H_{\eta_L}}{\partial p^2}(x, t\, d_x u)(d_x u)^2\, dt.$$

Let us observe the following.

- From the previous lemma, one has that
$$\frac{\partial H_{\eta_L}}{\partial p}(x, 0) = 0,$$
for every $x \in M$.
- From the strict convexity hypothesis, it follows that there exists $\gamma > 0$ such that
$$\frac{\partial^2 H}{\partial p^2}(x, t\, d_x u)(d_x u)^2 \geq \gamma \|d_x u\|_x^2$$
for all $x \in M$ and $0 \leq t \leq 1$.

Therefore, for $x \in W$,

$$\alpha(c_L) \geq H_{\eta_L}(x, d_x u) \geq H_{\eta_L}(x, 0) + \frac{\gamma}{2}\|d_x u\|_x^2 = -L(x, 0) + \frac{\gamma}{2}\|d_x u\|_x^2.$$

The assertion will follow from the previous lemma, choosing

$$U(x) = \frac{2}{\gamma}(\alpha(c_L) + L(x, 0)).$$

In fact, $U \in C^r$, with $r \geq 2d - 2$, by hypothesis; moreover, it satisfies all other conditions, because

$$\alpha(c_L) = -\inf_{x \in M} L(x, 0)$$

and

$$\mathcal{A}_{L,c_L} \subseteq \{x \in W : L(x, 0) = -\alpha(c_L)\} = \{x \in W : U(x) = 0\} =: \mathcal{A}.$$

So, the previous lemma allows us to conclude that

$$|u(\mathcal{A}_{L,c_L})| = 0.$$

□

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

Proof of Main theorem. Suppose by contradiction that $\bar{\mathcal{A}}_{L,c_L}$ is not totally disconnected; therefore, it must contain a connected component $\overline{\Gamma}$ with at least two points $\bar{x}$ and $\bar{y}$. In particular,

$$\bar{\delta}_c(\bar{x}, \bar{y}) = h_{\eta_L}(x, y) + h_{\eta_L}(y, x) > 0,$$

for some $x \in \bar{x}$ and $y \in \bar{y}$; therefore, we have $h_{\eta_L}(x, y) > 0$ or $h_{\eta_L}(y, x) > 0$. From the representation formula for h_{η_L}, it follows that there exists $u \in \mathcal{S}^{1,1}_{\eta_L} \subseteq \mathcal{D}^{1,1}_{c_L}$ (since $u = u - 0$, and $v = 0$ is a η_L-critical subsolution), such that $|u(y) - u(x)| > 0$.

This implies that the set $\varphi_u(\overline{\Gamma})$ is a connected set in $\mathbb{R}$ with at least two different points, and hence it is a non-degenerate interval and its Lebesgue measure is positive. But

$$\varphi_u(\overline{\Gamma}) \subseteq \varphi_u(\bar{\mathcal{A}}_{L,c_L}) = u(\mathcal{A}_{L,c_L})$$

and consequently

$$0 < |\varphi_u(\overline{\Gamma})| \leq |u(\mathcal{A}_{L,c_L})|.$$

This contradicts the previous corollary. □

In particular, this proof suggests a possible approach to generalize the above result to more general Lagrangians and other cohomology classes.

Definition. A C^1 function $f : M \longrightarrow \mathbb{R}$ is of *Morse–Sard type* if $|f(\mathrm{Crit}(f))| = 0$, where $\mathrm{Crit}(f)$ is the set of critical points of f and $|\cdot|$ denotes the Lebesgue measure in $\mathbb{R}$.

PROPOSITION 6. *Let M be a compact connected manifold of dimension $d \geq 1$, L a Tonelli Lagrangian and $c \in H^1(M; \mathbb{R})$. If each $w \in \mathcal{D}^{1,1}_c$ is of Morse–Sard type, then the quotient Aubry set $(\bar{\mathcal{A}}_{L,c}, \bar{\delta}_c)$ is totally disconnected.*

This proposition and Sard's lemma (see [**3**]) easily imply Mather's result in dimension $d \leq 2$ (autonomous case); it suffices to note that Sard's lemma (in dimension d) holds for $C^{d-1,1}$ functions.

COROLLARY 6. *Let M be a compact connected manifold of dimension $d \leq 2$. For any L Tonelli Lagrangian and $c \in H^1(M; \mathbb{R})$, the quotient Aubry set $(\bar{\mathcal{A}}_{L,c}, \bar{\delta}_c)$ is totally disconnected.*

Remark. The main problem becomes now to understand under which conditions on L and c these differences of subsolutions are of *Morse–Sard type*. Unfortunately, one cannot use the classical Sard's lemma, due to a lack of regularity of critical subsolutions: in general, they will be at most $C^{1,1}$. In fact, although it is always possible to smooth them up out of the Aubry set and obtain functions in $C^\infty(M \setminus \mathcal{A}_{L,c}) \cap C^{1,1}(M)$, the presence of the Aubry set (where the value of their differential is prescribed) represents an obstacle that it is impossible to overcome. It is quite easy to construct examples that do not admit C^2 critical subsolutions: just consider a case in which $\mathcal{A}_{L,c}$ is all the manifold and it is not a C^1 graph. For instance, this is the case if $M = \mathbb{T}$ and $H(x, p) = \frac{1}{2}(p + (2/\pi))^2 - \sin^2(\pi x)$; in fact, there is only one critical subsolution (up to constants) that turns out to be a solution ($\mathcal{A}_{L,2/\pi} = \mathbb{T}$), and it is given by a primitive of $\sin(\pi x) - (2/\pi)$; this is clearly $C^{1,1}$ but not C^2.

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

On the other hand, the above results suggest that, in order to prove the Morse–Sard property, one could try to control the *complexity* of these functions (*à la* Yomdin), using the rigid structure provided by the Hamilton–Jacobi equation and the smoothness of the Hamiltonian, rather than the regularity of the subsolutions. There are several difficulties in pursuing this approach in the general case, mostly related to the nature of the Aubry set. We hope to understand these 'speculations' in more depth in the future.

4. *Proof of the Main Lemma*

Definition. Consider a function $f \in C^r(\mathbb{R}^d)$. We say that f is *s-flat* at $x_0 \in \mathbb{R}^d$ (with $s \leq r$), if all its derivatives, up to the order s, vanish at x_0.

The proof of the Main Lemma is based on the following version of *Kneser–Glaeser's Rough composition theorem* (see [**1, 18**]).

PROPOSITION 7. *Let V, $W \subset \mathbb{R}^d$ be open sets and $A \subset V$, $A^* \subset W$ closed sets. Consider $U \in C^r(V)$, with $r \geq 2$, a non-negative function that is s-flat on $A \subset \{U(x) = 0\}$, with $s \leq r - 1$, and $g : W \longrightarrow V$ a C^{r-s} function, with $g(A^*) \subset A$.*

Then, for every open pre-compact set $W_1 \supset A^*$ properly contained in W, there exists

$$F : \mathbb{R}^d \longrightarrow \mathbb{R}$$

satisfying the following properties:
(i) $F \in C^{r-1}(\mathbb{R}^d)$;
(ii) $F \geq 0$;
(iii) $F(x) = U(g(x)) = 0$ *on* A^*;
(iv) F *is s-flat on* A^*;
(v) $\{F(x) = 0\} \cap W_1 = A^*$;
(vi) *there exists a constant $K > 0$, such that $U(g(x)) \leq K F(x)$ on W_1.*

See §5 for its proof.

To prove the Main Lemma, it will be enough to show that, for every $x_0 \in M$, there exists a neighborhood Ω such that it holds. For such a local result, we can assume that $M = \mathcal{U}$ is an open subset of $\mathbb{R}^d$, with $x_0 \in \mathcal{U}$. In the following, we shall identify $\mathrm{T}^*\mathcal{U}$ with $\mathcal{U} \times \mathbb{R}^d$ and, for $x \in \mathcal{U}$, we identify $\mathrm{T}^*_x\mathcal{U} = \{x\} \times \mathbb{R}^d$. We equip $\mathcal{U} \times \mathbb{R}^d$ with the natural coordinates $(x_1, \ldots, x_d, p_1, \ldots, p_d)$.

Before proceeding with the proof, let us point out that it is locally possible to replace the norm obtained by the Riemannian metric by a constant norm on $\mathbb{R}^d$.

LEMMA 2. *For each $0 < \alpha < 1$ and $x_0 \in M$, there exists an open neighborhood Ω of x_0, with $\overline{\Omega} \subset \mathcal{U}$ and such that*

$$(1 - \alpha)\|p\|_{x_0} \leq \|p\|_x \leq (1 + \alpha)\|p\|_{x_0},$$

for every $p \in \mathrm{T}^*_x\mathcal{U} \cong \mathbb{R}^d$ and each $x \in \overline{\Omega}$.

Proof. By continuity of the Riemannian metric, the norm $\|p\|_x$ tends uniformly to 1 on $\{p : \|p\|_{x_0} = 1\}$, as x tends to x_0. Therefore, for x near to x_0 and every $p \in \mathbb{R}^d \setminus \{0\}$,

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

we have

$$(1-\alpha) \leq \left\| \frac{p}{\|p\|_{x_0}} \right\|_x \leq (1+\alpha). \qquad \square$$

We can now prove the main result of this section.

Proof of the Main Lemma. By choosing local charts and by Lemma 2, we can assume that $U \in C^r(\Omega)$, with Ω an open set in $\mathbb{R}^d$, $\mathcal{A} = \{x \in \Omega :\ U(x) = 0\}$ and $u : \Omega \longrightarrow \mathbb{R}$ such that $\|d_x u\|^2 \leq \beta U(x)$ in Ω, where β is a positive constant.

Define, for $1 \leq s \leq r$,

$$B_s = \{x \in \mathcal{A} :\ U \text{ is } s\text{-flat at } x\}$$

and observe that

$$\mathcal{A} = B_1 := \{x \in \mathcal{A} : DU(x) = 0\}.$$

We shall prove the lemma by induction on the dimension d. Let us start with the following claim.

CLAIM. *If $s \geq 2d - 2$, then $|u(B_s)| = 0$.*

Proof. Let $C \subset \Omega$ be a closed cube with edges parallel to the coordinate axes. We shall show that $|u(B_s \cap C)| = 0$. Since B_s can be covered by countably many such cubes, this will prove that $|u(B_s)| = 0$.

Let us start observing that, by Taylor's theorem, for any $x \in B_s \cap C$ and $y \in C$ we have

$$U(y) = R_s(x;\, y),$$

where $R_s(x;\, y)$ is Taylor's remainder. Therefore, for any $y \in C$,

$$U(y) = o(\|y - x\|^s).$$

Let λ be the length of the edge of C. Choose an integer $N > 0$ and subdivide C into N^d cubes C_i with edges λ/N, and order them so that, for $1 \leq i \leq N_0 \leq N^d$, one has $C_i \cap B_s \neq \emptyset$. Hence,

$$B_s \cap C = \bigcup_{i=1}^{N_0} B_s \cap C_i.$$

Observe that, for every $\varepsilon > 0$, there exists $\nu_0 = \nu_0(\varepsilon)$ such that, if $N \geq \nu_0$, $x \in B_s \cap C_i$ and $y \in C_i$, for some $0 \leq i \leq N_0$, then

$$U(y) \leq \frac{\varepsilon^2}{4\beta(d\lambda^2)^d} \|y - x\|^s.$$

Fix $\varepsilon > 0$. Choose $x_i \in B_s \cap C_i$ and call $y_i = u(x_i)$. Define, for $N \geq \nu_0$, the following intervals in $\mathbb{R}$:

$$E_i = \left[y_i - \frac{\varepsilon}{2N^d},\ y_i + \frac{\varepsilon}{2N^d} \right].$$

Let us show that, if N is sufficiently large, then $u(B_s \cap C) \subset \bigcup_{i=1}^{N_0} E_i$.

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

In fact, if $x \in B_s \cap C$, then there exists $1 \le i \le N_0$, such that $x \in B_s \cap C_i$. Therefore,

$$\begin{aligned}|u(x) - y_i| &= |u(x) - u(x_i)| \\ &= \|d_x u(\tilde{x})\| \cdot \|x - x_i\| \\ &\le \sqrt{\beta U(\tilde{x})}\|x - x_i\| \\ &\le \sqrt{\beta \frac{\varepsilon^2}{4\beta(d\lambda^2)^d}\|\tilde{x} - x_i\|^{s/2}}\|x - x_i\| \\ &\le \frac{\varepsilon}{2(d\lambda^2)^{d/2}}\|x - x_i\|^{(s+2)/2} \\ &\le \frac{\varepsilon}{2(d\lambda^2)^{d/2}}\left(\sqrt{d}\frac{\lambda}{N}\right)^{(s+2)/2},\end{aligned}$$

where $\tilde{x}$ is a point in the segment joining x and x_i. Since by hypothesis $s \ge 2d - 2$, then $(s + 2)/2 \ge d$. Hence, assuming that $N > \max\{\lambda\sqrt{d}, \nu_0\}$, one gets

$$|u(x) - y_i| \le \frac{\varepsilon}{2N^d}$$

and one can deduce the inclusion above.

To prove the claim, it is now enough to observe that

$$\begin{aligned}|u(B_s \cap C)| &\le \left|\bigcup_{i=1}^{N_0} E_i\right| \le \sum_{i=1}^{N_0} |E_i| \\ &\le \varepsilon N_0 \frac{1}{N^d} \\ &\le \varepsilon N^d \frac{1}{N^d} \\ &= \varepsilon.\end{aligned}$$

From the arbitrariness of ε, the assertion follows easily. □

This claim immediately implies that $u(B_{2d-2})$ has measure zero.

In particular, this proves the case $d = 1$ (since in this case $2d - 2 = 0$) and it allows us to start the induction.

Suppose that we have proven the result for $d - 1$ and we want to show it for d. Since

$$\mathcal{A} = (B_1 \setminus B_2) \cup (B_2 \setminus B_3) \cup \cdots \cup (B_{2d-3} \setminus B_{2d-2}) \cup B_{2d-2},$$

it remains to show that $|u(B_s \setminus B_{s+1})| = 0$ for $1 \le s \le 2d - 3 \le r - 1$.

CLAIM. *Every $\tilde{x} \in B_s \setminus B_{s+1}$ has a neighborhood $\tilde{V}$, such that*

$$|u((B_s \setminus B_{s+1}) \cap \tilde{V})| = 0.$$

Since $B_s \setminus B_{s+1}$ can be covered by countably many such neighborhoods, this implies that $u(B_s \setminus B_{s+1})$ has measure zero.

Proof. Choose $\tilde{x} \in B_s \setminus B_{s+1}$. By definition of these sets, all partial derivatives of order s of U vanish at this point, but there is one of order $s + 1$ that does not. Assume (without any loss of generality) that there exists a function

$$w(x) = \partial_{i_1}\partial_{i_2} \cdots \partial_{i_s} U(x)$$

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

such that

$$w(\tilde{x}) = 0 \quad \text{but } \partial_1 w(\tilde{x}) \neq 0.$$

Define

$$\begin{aligned} h : \Omega &\longrightarrow \mathbb{R}^d \\ x &\longmapsto (w(x), x_2, \ldots, x_d), \end{aligned}$$

where $x = (x_1, x_2, \ldots, x_d)$. Clearly, $h \in C^{r-s}(\Omega)$ and $Dh(\tilde{x})$ is non-singular; hence, there is an open neighborhood V of $\tilde{x}$ such that

$$h : V \longrightarrow W$$

is a C^{r-s} isomorphism (with $W = h(V)$).

Let V_1 be an open precompact set, containing $\tilde{x}$ and properly contained in V, and define $A = B_s \cap \overline{V_1}$, $A^* = h(A)$ and $g = h^{-1}$. If we consider W_1, any open set containing A^* and properly contained in W, we can apply Proposition 7 and deduce the existence of $F : \mathbb{R}^d \longrightarrow \mathbb{R}$ satisfying properties (i)–(vi).

Define

$$\hat{W} = \{(x_2, \ldots, x_d) \in \mathbb{R}^{d-1} : (0, x_2, \ldots, x_d) \in W_1\}$$

and

$$\hat{U}(x_2, \ldots, x_d) = C F(0, x_2, \ldots, x_d),$$

where C is a positive constant to be chosen sufficiently large. Observe that $\hat{U} \in C^{r-1}(\mathbb{R}^{d-1})$.

Moreover, property (v) of F and the fact that $A^* = h(A) \subseteq \{0\} \times \hat{W}$ imply that

$$A^* = \{0\} \times \hat{B}_1,$$

where $\hat{B}_1 = \{(x_2, \ldots, x_d) \in \hat{W} : \ F(0, x_2, \ldots, x_d) = 0\}$. Denote

$$\hat{\mathcal{A}} := \{(x_2, \ldots, x_d) \in \hat{W} : \ \hat{U} = 0\} = \hat{B}_1$$

and define the following function on $\hat{W}$:

$$\hat{u}(x_2, \ldots, x_d) = u(g(0, x_2, \ldots, x_d)).$$

We want to show that these functions satisfy the hypotheses for the $(d-1)$-dimensional case. In fact:

- $\hat{U} \in C^{r-1}(\mathbb{R}^{d-1})$, with $r - 1 \geq 2d - 3 > 2(d-1) - 2$;
- $\hat{u} \in C^1(\hat{W})$ (since g is in $C^{r-s}(W)$, where $1 \leq s \leq r - 1$);
- if we denote $\mu = \sup_{W_1} \|d_x g\| < +\infty$ (since g is C^1 on $\overline{W_1}$), then we have that for every point in $\hat{W}$,

$$\begin{aligned} \|d\hat{u}(x_2, \ldots, x_d)\|^2 &\leq \|d_x u(g(0, x_2, \ldots, x_d))\|^2 \|d_x g(0, x_2, \ldots, x_d)\|^2 \\ &\leq \mu^2 \|d_x u(g(0, x_2, \ldots, x_d))\|^2 \\ &\leq \beta\mu^2 U(g(0, x_2, \ldots, x_d)) \\ &\leq \beta\mu^2 K F(0, x_2, \ldots, x_d) \\ &\leq \hat{U}(x_2, \ldots, x_d), \end{aligned}$$

  if we choose $C > \beta\mu^2 K$, where K is the positive constant appearing in Proposition 7, property (vi).

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

Therefore, it follows from the inductive hypothesis, that

$$|\hat{u}(\hat{\mathcal{A}})| = 0.$$

Since,

$$\begin{aligned} u(B_s \cap V_1) \subseteq u(A) = u(g(A^*)) = u(g(\{0\} \times \hat{B}_1)) \\ = \hat{u}(\hat{B}_1) = \hat{u}(\hat{\mathcal{A}}), \end{aligned}$$

defining $\tilde{V} = V_1$, we may conclude that

$$|u(B_s \cap \tilde{V})| \leq |\hat{u}(\hat{\mathcal{A}})| = 0. \qquad \square$$

This completes the proof of the Main Lemma. □

5. *Proof of a modified version of Kneser–Glaeser's rough composition theorem*

Now, let us prove Proposition 7. We shall mainly follow the presentation in [**1**], adapted to our needs.

Proof of Proposition 7. Let us start by defining a family of polynomials. Supposing that g is C^r and using the s-flatness hypothesis, we have, for $x \in A^*$ and $k = 0, 1, \ldots, r$,

$$f_k(x) = D^k(U \circ g)(x) = \sum_{s<q\leq k} \sum \sigma_k D^q U(g(x)) D^{i_1} g(x) \cdots D^{i_q} g(x), \tag{1}$$

where the second sum is over all the q-tuples of integers $i_1, \ldots, i_q \geq 1$ such that $i_1 + \cdots + i_q = k$, and $\sigma_k = \sigma_k(i_1, \ldots, i_q)$.

The crucial observation is that (1) makes sense on A^*, even when g is C^{r-s} smooth (in fact, $i_j \leq k - q + 1 \leq r - s$).

We would like to proceed in the fashion of *Whitney's extension theorem*, in order to find a smooth function F such that $D^k F = f_k$ on A^*, and satisfying the stated conditions.

Remark. Note that, without any loss of generality, we can assume that W is contained in an open ball of diameter one. The general case will then follow from this special one, by a straightforward partition of unity argument.

Let us start with some technical lemmata.

LEMMA 3. *For $x, x', x_0 \in A^*$ and $k = 0, \ldots, r$, we have*

$$f_k(x') = \sum_{i\leq r-k} \frac{f_{k+i}(x)}{i!}(x' - x)^i + R_k(x, x'),$$

with

$$\frac{|R_k(x, x')|}{\|x' - x\|^{r-k}} \longrightarrow 0$$

as $x, x' \longrightarrow x_0$ in A^*.

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

The proof of this lemma appears without any major modification in [**1,** pp. 36–37].

Define, for $x \in A^*$ and $y \in \mathbb{R}^d$,

$$P(x, y) = \sum_{i=s+1}^{r} \frac{f_i(x)}{i!}(y-x)^i$$

and its kth derivative

$$P_k(x, y) = \sum_{i \le r-k} \frac{f_{i+k}(x)}{i!}(y-x)^i.$$

LEMMA 4. *For $x \in A^*$ and $y \in W_1$,*

$$U(g(y)) = P(x, y) + R(x, y),$$

where $|R(x, y)| \le C\|y-x\|^r$.

Proof. The proof follows the same idea of Lemma 3. By Taylor's formula for U,

$$U(g(y)) = \sum_{q=s+1}^{r} \frac{D^q U(g(x))}{q!}(g(y)-g(x))^q + I(g(x), g(y))(g(x)-g(y))^r.$$

Obviously,

$$|I(g(x), g(y))(g(x)-g(y))^r| \le C_1\|y-x\|^r,$$

so therefore it is sufficient to estimate the first term.

Observe that

$$g(y) = g(x) + \sum_{i=1}^{r-s} D^i g(x)(y-x)^i + J(x, y)(y-x)^{r-s}.$$

Hence, the first term in the sum above becomes

$$\begin{aligned}\sum_{q=s+1}^{r} \frac{D^q U(g(x))}{q!}&\left[\sum_{i=1}^{r-s} D^i g(x)(y-x)^i + J(x, y)(y-x)^{r-s}\right]^q \\ &= \sum_{k=s+1}^{r} a_k (y-x)^k + \hat{R}(x, y) \\ &= P(x, y) + \hat{R}(x, y),\end{aligned}$$

since

$$a_k = \sum_{s+1 \le q \le k} \sum D^q U(g(x)) D^{i_1} g(x) \cdots D^{i_q} g(x) = \frac{f_k(x)}{k!}.$$

The remainder of the terms consist of:

- terms containing $(y-x)^k$, with $k > r$;
- terms of the binomial product, containing $J(x, y)(y-x)^{r-s}$. They are of the form

$$\ldots (y-x)^{(r-s)j + \sum_{i=1}^{r-s} i\alpha_i}$$

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

where $\alpha_i \geq 0$ and $\sum \alpha_i = q - j$. Since $q \geq s + 1$ and $s \leq r - 1$, then

$$\begin{aligned}(r-s)j + \sum_{i=1}^{r-s} i\alpha_i &\geq (r-s)j + \sum_{i=1}^{r-s} \alpha_i \\ &= (r-s)j + q - j \\ &= rj - sj + q - j \\ &\geq rj - (s+1)j + s + 1 \\ &= r + r(j-1) - (s+1)(j-1) \\ &= r + (r-s-1)(j-1) \geq r.\end{aligned}$$

Therefore, for $x \in A^*$ and $y \in W_1$,

$$|\hat{R}(x, y)| \leq C_2 \|y - x\|^r,$$

and the lemma follows taking $C = C_1 + C_2$. □

The next step will consist of creating a *Whitney's partition*. We will start by covering $W_1 \setminus A^*$ with an infinite collection of cubes K_j, such that the size of each K_j is roughly proportional to its distance from A^*.

First, let us fix some notation. We shall write $a \prec b$ instead of 'there exists a positive real constant M, such that $a \leq Mb$' and $a \approx b$ as short for $a \prec b$ and $b \prec a$.

Let $\lambda = 1/(4\sqrt{d})$; this choice will come in handy later. For any closed cube K (with edges parallel to the coordinate axes), K^λ will denote the $(1 + \lambda)$-dilation of K about its center.

Let $\|\cdot\|$ be the euclidean metric on $\mathbb{R}^d$ and

$$d(y) = d(y, A^*) = \inf\{\|y - x\| : x \in A^*\}.$$

If $\{K_j\}_j$ is the sequence of closed cubes defined below, with edges of length e_j, let d_j be its distance from A^*, i.e.

$$d_j = d(A^*, K_j) = \inf\{\|y - x\| : x \in A^*, \ y \in K_j\}.$$

One can show the following classical lemma (see, for instance, [**1**] for a proof).

LEMMA 5. *There exists a sequence of closed cubes $\{K_j\}_j$ with edges parallel to the coordinate axes that satisfies the following properties:*

(i) *the interiors of the K_j are disjoint;*

(ii) $W_1 \setminus A^* \subset \bigcup_j K_j$;

(iii) $e_j \approx d_j$;

(iv) $e_j \approx d(y)$ *for all* $y \in K_j^\lambda$;

(v) $e_j \approx d(z)$ *for all* $z \in W_1 \setminus A^*$, *such that the ball with center z and radius $\frac{1}{8}d(z)$ intersects K_j^λ;*

(vi) *each point of $W_1 \setminus A^*$ has a neighborhood intersecting at most N of the K_j^λ, where N is an integer depending only on d.*

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

Now, let us construct a partition of unity on $W_1 \setminus A^*$. Let Q be the unit cube centered at the origin. Let η be a C^∞ bump function defined on $\mathbb{R}^d$ such that

$$\eta(y) = \begin{cases} 1 & \text{for } y \in Q, \\ 0 & \text{for } y \notin Q^\lambda, \end{cases}$$

and $0 \leq \eta \leq 1$. Define

$$\eta_j(y) = \eta\left(\frac{y - c_j}{e_j}\right),$$

where c_j is the center of K_j and e_j is the length of its edge, and consider

$$\sigma(y) = \sum_j \eta_j(y).$$

Then, $1 \leq \sigma(y) \leq N$ for all $y \in W_1 \setminus A^*$. Clearly, for each $k = 0, 1, 2, \ldots$ we have that $D^k \eta_j(y) \prec e_j^{-k}$, for all $y \in W_1 \setminus A^*$. Hence, by properties (iv) and (vi) of Lemma 5, we have that, for each $k = 0, 1, \ldots, r$,

$$D^k \eta_j(y) \prec d(y)^{-k} \quad \text{for all } y \in W_1 \setminus A^*$$

and

$$D^k \sigma(y) \prec d(y)^{-k} \quad \text{for all } y \in W_1 \setminus A^*.$$

Define

$$\varphi_j(y) = \frac{\eta_j(y)}{\sigma(y)}.$$

These functions satisfy the following properties:
(i) each φ_j is C^∞ and supported on K_j^λ;
(ii) $0 \leq \varphi_j(y) \leq 1$ and $\sum_j \varphi_j(y) = 1$, for all $y \in W_1 \setminus A^*$;
(iii) every point of $W_1 \setminus A^*$ has a neighborhood on which all but at most N of the φ_j vanish identically;
(iv) for each $k = 0, 1, \ldots, r$, $D^k \varphi_j(y) \prec d(y)^{-k}$ for all $y \in W_1 \setminus A^*$; namely, there are constants M_k such that $D^k \varphi_j(y) \leq M_k d(y)^{-k}$;
(v) there is a constant α and points $x_j \in A^*$, such that

$$\|x_j - y\| \leq \alpha d(y) \quad \text{whenever } \varphi_j(y) \neq 0.$$

This follows from properties (iii) and (iv) of Lemma 5.

We can now construct our function F. Observe that, from Lemma 4,

$$0 \leq U(g(y)) = P(x_j, y) + R(x_j, y) \leq P(x_j, y) + C\|y - x_j\|^r;$$

therefore, $P(x_j, y) \geq -C\|y - x_j\|^r$.

First, define

$$\hat{P}_j(y) = P(x_j, y) + 2C\|y - x_j\|^r$$

where C is the same constant as in Lemma 4; from what is said above,

$$\hat{P}_j(y) \geq C\|y - x_j\|^r > 0 \quad \text{in } W_1 \setminus \{x_j\}. \tag{2}$$

Hence, construct F in the following way:

$$F(y) = \begin{cases} 0, & y \in A^*, \\ \sum_j \varphi_j(y)\hat{P}_j(y), & y \in \mathbb{R}^d \setminus A^*. \end{cases}$$

We claim that this satisfies all the stated properties (i)–(vi). In particular, properties (ii), (iii) and (v) follow immediately from the definition of F and (2). Moreover, $F \in C^\infty(\mathbb{R}^d \setminus A^*)$. We need to show that $D^k F = f_k$ (for $k = 0, 1, \dots, r-1$) on ∂A^* (namely, the boundary of A^*) and that $D^{r-1} F$ is continuous on it. The main difficult in the proof is that $D^k F$ is expressed as a sum containing terms

$$D^{k-m}\varphi_j(y) P_m(x_j, y),$$

where $\varphi_j(y) \neq 0$. Even if y is close to some $x_0 \in A^*$, it could be closer to A^* and hence the bound given by property (iv) of φ_j might become large. One can overcome this problem by choosing a point $x^* \in A^*$, so that $\|x^* - y\|$ is roughly the same as $d(y)$ and hence x_j is close to x^*.

LEMMA 6. *For every $\eta > 0$, there exists $\delta > 0$ such that, for all $y \in W_1 \setminus A^*$, $x, x^* \in A^*$ and $x_0 \in \partial A^*$, we have*

$$\|P_k(x, y) - P_k(x^*, y)\| \le \eta\, d(y)^{r-k} \le \eta \|y - x_0\|^{r-k},$$

whenever $k \le r$ and

$$\begin{cases} \|y - x\| < \alpha\, d(y) \\ \|y - x^*\| < \alpha\, d(y) \\ \|y - x_0\| < \delta, \end{cases}$$

where α is the same constant as in (v) *above.*

See [**1,** p. 126] for its proof.

LEMMA 7. *For every $\eta > 0$, there exist $0 < \delta < 1$ and a constant E such that, for all $y \in W_1 \setminus A^*$, $x^* \in A^*$ and $x_0 \in \partial A^*$, we have*

$$\|D^k F(y) - P_k(x^*, y)\| \le E\, d(y)^{r-k} \le \eta\, d(y)^{r-k-1},$$

whenever $k \le r - 1$ and

$$\begin{cases} \|y - x^*\| < \alpha\, d(y) \\ \|y - x_0\| < \delta. \end{cases}$$

Proof. Let

$$S_{j,k}(x^*, y) = \partial_k \hat{P}_j(y) - P_k(x^*, y).$$

From Lemma 6 (with $\eta = \varepsilon$, to be defined later) and the definition of $\hat{P}_j$, we get

$$\begin{aligned} \|S_{j,k}(x^*, y)\| &\le \|\partial_k \hat{P}_j(y) - P_k(x_j, y)\| + \|P_k(x_j, y) - P_k(x^*, y)\| \\ &\le C_k d(y)^{r-k} + \varepsilon\, d(y)^{r-k} \\ &= (C_k + \varepsilon)\, d(y)^{r-k}. \end{aligned}$$

Then,

$$F(y) - P(x^*, y) = \sum_j \varphi_j(y) S_{j,0}(x^*, y)$$

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

and hence

$$D^k F(y) - P_k(x^*, y) = \sum_j \sum_{i \le k} \binom{k}{i} D^{k-i} \varphi_j(y) S_{j,i}(x^*, y).$$

Therefore, choosing ε sufficiently small,

$$\begin{aligned}\|D^k F(y) - P_k(x^*, y)\| &\le \sum_j \sum_{i \le k} \binom{k}{i} \|D^{k-i}\varphi_j(y)\| \cdot \|S_{j,i}(x^*, y)\| \\ &\le \sum_j \sum_{i \le k} \binom{k}{i} M_{k-i}\, d(y)^{-k+i} (C_k + \varepsilon)\, d(y)^{r-i} \\ &\le E\, d(y)^{r-k} \le \eta\, d(y)^{r-k-1}.\end{aligned}$$ □

□

LEMMA 8. *For every* $\eta > 0$, *there exists* $0 < \delta < 1$ *such that, for all* $y \in W_1 \setminus A^*$, $x^* \in A^*$ *and* $x_0 \in \partial A^*$, *we have*

$$\|P_k(x^*, y) - P_k(x_0, y)\| \le \eta \|y - x_0\|^{r-k},$$

whenever $k \le r$ *and*

$$\begin{cases} \|y - x^*\| < \alpha\, d(y) \\ \|y - x_0\| < \delta. \end{cases}$$

Proof. The proof goes as in the proof of Lemma 6, observing that $\|x^* - x_0\| \le (1 + \alpha)\|y - x_0\|$ and

$$P_k(x_0, y) - P_k(x^*, y) = \sum_{q \le r-k} \frac{R_{k+q}(x^*, x_0)}{q!} (y - x)^q.$$ □

CLAIM. *For every* $x_0 \in \partial A^*$ *and* $k = 0, 1, \ldots, r - 1$,

$$D^k F(x_0) = f_k(x_0).$$

Moreover, $D^{r-1}F$ *is continuous at* $x_0 \in \partial A^*$.

This claim follows easily from the lemmata above (see [**1,** p. 128] for more details).

This proves that $F \in C^{r-1}(\mathbb{R}^d)$ and completes the proof of (i) and (iv).

It remains to show that property (vi) holds, namely, that there exists a constant $K > 0$, such that $U(g(x)) \le K F(x)$ on W_1. Obviously, this holds at every point in A^*, for every choice of K (since both functions vanish there).

CLAIM. *There exists a constant* $K > 0$, *such that* $(U \circ g)/F \le K$ *on* $W_1 \setminus A^*$.

Proof. Since $F > 0$ on $W_1 \setminus A^*$, it is sufficient to show that $(U \circ g)/F$ is uniformly bounded by a constant, as $d(y)$ goes to zero.

Let us start by observing that, for $y \in K_j^\lambda$,

$$\hat{P}_j(y) \ge C \|y - x_j\|^r \ge C\, d(y)^r;$$

https://doi.org/10.1017/S0143385707000351 Published online by Cambridge University Press

therefore,

$$\begin{aligned} F(y) &= \sum_j \varphi_j(y)\hat{P}_j(y) \\ &\geq \sum_j \varphi_j(y) C\, d(y)^r \\ &= C\, d(y)^r. \end{aligned}$$

Moreover, if $x^* \in A^*$ such that $d(y) = \|y - x^*\|$, Lemmas 4 and 7 imply that

$$\begin{aligned} |U(g(y)) - F(y)| &\leq |U(g(y)) - P(x^*, y)| + |P(x^*, y) - F(y)| \\ &\leq C\, d(y)^r + E\, d(y)^r = (C + E)\, d(y)^r. \end{aligned}$$

Hence,

$$\begin{aligned} \frac{U(g(y))}{F(y)} &= \frac{U(g(y)) - F(y) + F(y)}{F(y)} \\ &\leq 1 + \frac{|U(g(y)) - F(y)|}{F(y)} \\ &\leq 1 + \frac{(C + E)\, d(y)^r}{C\, d(y)^r} \\ &\leq 2 + \frac{E}{C} =: \tilde{K}. \end{aligned}$$ □

This proves property (vi) and concludes the proof of the proposition. □

Acknowledgements. I wish to thank John Mather for having introduced me to this area and suggesting this problem. I am very grateful to him and to Albert Fathi for their interest and for several helpful discussions.

REFERENCES

[1] R. Abraham and J. Robbin. *Transversal Mappings and Flows*. W. A. Benjamin, New York, 1967, With an appendix by Al Kelley.

[2] V. Bangert. Mather sets for twist maps and geodesics on tori. *Dynamics Reported, Vol. 1 (Dynam. Report. Ser. Dynam. Systems Appl., 1)*. Wiley, Chichester, 1988, pp. 1–56.

[3] S. M. Bates. Toward a precise smoothness hypothesis in Sard's theorem. *Proc. Amer. Math. Soc.* **117**(1) (1993), 279–283.

[4] P. Bernard. Existence of $C^{1,1}$ critical sub-solutions of the Hamilton–Jacobi equation on compact manifolds. *Ann. Sci. École Norm. Sup.* (4), to appear.

[5] D. Burago, S. Ivanov and B. Kleiner. On the structure of the stable norm of periodic metrics. *Math. Res. Lett.* **4**(6) (1997), 791–808.

[6] G. Contreras, J. Delgado and R. Iturriaga. Lagrangian flows: the dynamics of globally minimizing orbits. II. *Bol. Soc. Brasil. Mat. (N.S.)* **28**(2) (1997), 155–196.

[7] G. Contreras and R. Iturriaga. Global minimizers of autonomous Lagrangians. 22° *Colóquio Brasileiro de Matemática (22nd Brazilian Mathematics Colloquium)*. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.

[8] A. Fathi. Sard, Whitney, Assouad and Mather. *Talk at Recent and Future developments in Hamiltonian Systems* (*Institut Henri Poincaré, Paris, France, May 2005*).

[9] A. Fathi. *Weak KAM Theorem and Lagrangian Dynamics*. Cambridge University Press, Cambridge, to appear.

[10] A. Fathi and A. Siconolfi. Existence of C^1 critical subsolutions of the Hamilton–Jacobi equation. *Invent. Math.* **155**(2) (2004), 363–388.

[11] G. Forni and J. N. Mather. Action minimizing orbits in Hamiltonian systems. *Transition to Chaos in Classical and Quantum Mechanics (Montecatini Terme, 1991) (Lecture Notes in Mathematics, 1589)*. Springer, Berlin, 1994, pp. 92–186.

[12] P. Hajlasz. Whitney's example by way of Assouad's embedding. *Proc. Amer. Math. Soc.* **131**(11) (2003), 3463–3467 (electronic).

[13] R. Mañé. Lagrangian flows: the dynamics of globally minimizing orbits. *Bol. Soc. Brasil. Mat. (N.S.)* **28**(2) (1997), 141–153.

[14] J. N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.* **207**(2) (1991), 169–207.

[15] J. N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)* **43**(5) (1993), 1349–1386.

[16] J. N. Mather. Total disconnectedness of the quotient Aubry set in low dimensions. *Comm. Pure Appl. Math.* **56**(8) (2003), 1178–1183. Dedicated to the memory of Jürgen K. Moser.

[17] J. N. Mather. Examples of Aubry sets. *Ergod. Th. & Dynam. Sys.* **24**(5) (2004), 1667–1723.

[18] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36**(1) (1934), 63–89.