Ergod. Th. & Dynam. Sys. (2004), **24**, 1667–1723
DOI: 10.1017/S0143385704000446
© *2004 Cambridge University Press*
Printed in the United Kingdom

Examples of Aubry sets

JOHN N. MATHER

Princeton University, Department of Mathematics, Princeton, NJ 08544-1000, USA
(e-mail: jnm@math.princeton.edu)

(*Received* 16 *December* 2003 *and accepted in revised form* 30 *June* 2004)

In memory of Michel Herman

Abstract. We construct a quotient Aubry set isometric to the unit interval. We also construct several related examples of Aubry sets.

Contents

1 Aubry sets 1669
2 The main result 1672
3 Construction of Z 1674
4 A property of Z 1679
5 Construction of U 1682
6 The case $L = K + U$ and $c = 0$ 1686
7 A description of $\rho_{L_0,0}$ when $c_m \geq \nu_m^* b_{m+1}$ 1687
8 Proof of the Proposition in §7: first step 1689
9 Proof of the Proposition in §7: second step 1691
10 Proof of the Proposition in §7: third step 1694
11 Proof of the Proposition in §7: fourth step 1698
12 Proof of the Proposition in §7: fifth step 1702
13 End of the proof of the Proposition in §7 1705
14 End of the proof of the Proposition in §5 1707
15 Proof of Theorem 2 1708
16 Some connecting orbits 1709
17 An example where $\dot{M}_{L,c} = \dot{A}_{L,c} = \dot{N}_{L,c}$ and $\bar{A}_{L,c}$ is not reduced to one point 1717
18 The failure of upper semi-continuity of the Aubry set 1720
19 Total disconnectedness of $A_{L_0,0}$ when U is of class C^{2d-2} 1722
20 Open problems 1722

In [**Mat1, Mat2**], we began a study of the dynamics of a class of Lagrangian systems, using variational methods, especially minimization. In subsequent studies (for the most part not yet written), we have found that understanding certain aspects of the Aubry set seems to help to construct orbits with interesting properties by the method of minimization. The only one of these studies written to date is [**Mat5**], in which we announce a form of Arnold

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

diffusion in the time-periodic case in two degrees of freedom and in the autonomous case in three degrees of freedom. Our proof (not yet written) of these results involves a detailed study of the Aubry sets. In the cases relevant to our proof, the structure of the Aubry sets closely resembles the structure of the Aubry sets of twist maps, see e.g. [**B, FoM**] and references therein. In these cases, there is consequently a detailed structure theory of Aubry sets, and we use this theory in our proof.

In more degrees of freedom, there seems to be no such detailed structure theory of Aubry sets. An obvious strategy to generalize our proof to more degrees of freedom would be to single out a collection of properties of Aubry sets that are used in the proof and generalize these properties to more degrees of freedom. We will not do this here. We do not know exactly what properties are needed. However, it seems to us that, whatever ones are needed, they include the total disconnectedness of what we called in [**Mat4**] the *quotient Aubry set*.

Burago *et al* [**BIK**] have produced an example where a quotient Aubry set is not totally disconnected. Here, we produce further examples of quotient Aubry sets that are not totally disconnected. The main purpose of this paper is to construct an example of a Lagrangian L_0 having a quotient Aubry set that is isometric to a closed interval. The existence of such an example is stated as Theorem 1 in §2.

In §1, we define the Aubry set $A_{L,c}$ and the quotient Aubry set $\bar{A}_{L,c}$ associated to a Lagrangian and a cohomology class c.

Another purpose of this paper is to produce an example where the Aubry set is not upper semi-continuous as a (set-valued) function of the Lagrangian. In §18, we will construct a modification L_1 of L_0 and show that $L \mapsto A_{L,c}$ is not upper semi-continuous at L_1.

Conversely, if L_2 is a Lagrangian and c is a cohomology class for which $\bar{A}_{L_2,c}$ has vanishing one-dimensional Hausdorff measure, then (L_2, c) is a point of upper semi-continuity of the set-valued function $(L, c) \mapsto A_{L,c}$. We will not show this in this paper.

We do not regard these examples as producing an unbreachable barrier to generalizing our results [**Mat5**] to arbitrarily many degrees of freedom. They do seem to produce difficulties for using our methods to prove the existence of suitable orbits in the case of a Lagrangian L such that there is some cohomology class c for which $\bar{A}_{L,c}$ is not totally disconnected. However, to generalize the results of [**Mat5**], we should seek an open, dense sub-set of a cusp-residual set of L for which Arnold diffusion is displayed. Thus, it is probably more relevant for our program to ask whether there is a dense set of L such that $\bar{A}_{L,c}$ is totally disconnected for every cohomology class c, rather than ask whether $\bar{A}_{L,c}$ is totally disconnected for every L and every c. We give an exact formulation of this problem in §20.

All our examples have finite differentiability. Let $C^{r,\alpha}$ denote the class of functions that are r-times continuously differentiable and whose rth (total) derivative satisfies a Hölder condition of order α. Theorem 1 in §2 states that, for every $d \geq 2$ and every $\epsilon > 0$, there is a Lagrangian L whose differentiability class is $C^{2d-3,1-\epsilon}$ and whose state space is d-dimensional such that the quotient Aubry set $\bar{A}_{L,0}$ is isometric to the unit interval. Most of this paper is a proof of this result. We do not know whether there is a C^∞ Lagrangian L and a cohomology class c such that $\bar{A}_{L,c}$ is not totally disconnected. (See the problem in §2.)

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

1. *Aubry sets*

In several previous papers, we have studied the problem of constructing orbits in a special setting. In this section, for the convenience of the reader, we recall this setting and several notions associated to it. The purpose of this paper is to produce several non-obvious examples relating to these notions. Even though we have explained these notions in previous papers, it seems best to recall them, in order that the reader can see what the examples are about without needing to consult the previous papers.

The *state space* is a closed (i.e. compact, without boundary) differentiable manifold M. We consider a C^2 real-valued function defined on $TM \times \mathbb{T}/\mathbb{Z}$, where TM denotes the tangent bundle of M and $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the circle (or one-torus). We call such a function L a *Lagrangian* provided it satisfies a Legendre positive definiteness condition, superlinear growth in the fiber, and completeness of the associated Euler–Lagrange flow. An exact statement of these conditions is given in [**Mat1**].

In previous papers, we have defined several sub-sets of $TM \times \mathbb{T}$ and $M \times \mathbb{T}$, associated to a Lagrangian L and a cohomology class $c \in H^1(M; \mathbb{R})$. These are summarized by a commutative diagram:

$$\begin{array}{ccccccc} \dot{M}_{L,c} & \subset & \dot{A}_{L,c} & \subset & \dot{N}_{L,c} & \subset & TM \times \mathbb{T} \\ \downarrow & & \downarrow & & & & \downarrow \pi \\ M_{L,c} & \subset & A_{L,c} & & \subset & & M \times \mathbb{T}. \end{array}$$

Definitions of these sets may be found, for example, in [**Mat2**], where $\dot{N}_{L,c}$ is called $\dot{\Sigma}_c$.

The sets in the top row are invariant under the Euler–Lagrange flow and (except for $TM \times \mathbb{T}$) are defined by action minimizing principles.

Now we define these sets. For the purpose, we quote *verbatim* from [**Mat3**] and [**Mat4**].

If μ is an invariant probability measure on the *phase space* $TM \times \mathbb{T}$, then there are two quantities, the *average action* $A_L(\mu) = A(\mu) = A(\mu) \in \mathbb{R}$ and the *rotation vector* $\rho(\mu) \in H_1(M; \mathbb{R})$, which are associated to μ. (See [**Mat1**] or [**Mat2**] for detailed discussions.) The *minimal average action* is defined as

$$\beta_L(h) = \beta(h) = \min\{A(\mu) : \rho(\mu) = h\},$$

for $h \in H_1(M; \mathbb{R})$, where μ ranges over all Φ_L-invariant probability measures on phase space and Φ_L is the Euler–Lagrange flow associated to L. The function $\beta_L = \beta : H_1(M; \mathbb{R}) \longrightarrow \mathbb{R}$ is convex and has superlinear growth at infinity.

Its *convex conjugate* $\alpha_L = \alpha : H^1(M; \mathbb{R}) \longrightarrow \mathbb{R}$ is defined by $\alpha(c) = -\min_h\{\beta(h) - \langle h, c\rangle\}$, where h ranges over $H_1(M; \mathbb{R})$. This is also a convex function with superlinear growth.

If η is a 1-form on $M \times \mathbb{T}$, we let $\hat{\eta} : TM \times \mathbb{T} \longrightarrow \mathbb{R}$ be the function defined by $\hat{\eta}(v, \tau) = \langle v + 1_\tau, \eta(m, \tau)\rangle$, for $m \in M$, $v \in TM_m$, and $\tau \in \mathbb{T}$, where 1_τ denotes the positive unit vector in $T\mathbb{T}_\tau$ (often denoted $\partial/\partial\tau|_\tau$).

If η is a closed 1-form, we let $[\eta] = ([\eta]_M, [\eta]_{\mathbb{T}}) \in H^1(M \times \mathbb{T}; \mathbb{R}) = H^1(M; \mathbb{R}) \times \mathbb{R}$ denote its deRham cohomology class. We will say that η is *Mañé critical* if $[\eta]_{\mathbb{T}} = -\alpha([\eta]_M)$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Let μ be a probability measure on the phase space. If μ is invariant (for the Euler–Lagrange flow), then

$$\int (L - \hat{\eta})\, d\mu = A_L(\mu) - \langle \rho(\mu), [\eta]_M \rangle - [\eta]_{\mathbb{T}}.$$

If, in addition, η is Mañé critical, then

$$\min_{\mu} \left\{ \int (L - \hat{\eta})\, d\mu \right\} = 0,$$

where μ ranges over all Φ_L-invariant probability measures on phase space. In this case, we define $h_{L,\eta} = h_\eta$ by

$$h_\eta((m_0, \tau_0), (m_1, \tau_1)) = \lim_{T \to +\infty} \left\{ \inf_{\gamma} \int_a^b (L - \hat{\eta})(\gamma(t), \dot{\gamma}(t), t)\, dt \right\}$$

for $m_i \in M, \tau_i \in \mathbb{T} (i = 0, 1)$, where γ ranges over all absolutely continuous curves $\gamma : [a, b] \longrightarrow M$ such that $a \equiv \tau_0 \pmod 1$, $b \equiv \tau_1 \pmod 1$, $\gamma(a) = m_0$, $\gamma(b) = m_1$, and $b - a > T$. Setting $c = [\eta]_M$, we define $\rho_{L,c} = \rho_c$ by

$$\rho_c((m_0, \tau_0), (m_1, \tau_1)) = h_\eta((m_0, \tau_0), (m_1, \tau_1)) + h_\eta((m_1, \tau_1), (m_0, \tau_0)).$$

The functions $h_\eta, \rho_c : (M \times \mathbb{T})^2 \longrightarrow \mathbb{R}$ are essentially the same as functions defined in [**Mat2**]. (There, we considered only $\tau_0 \equiv \tau_1 \equiv 0 \pmod 1$, so these were functions from M^2 to $\mathbb{R}$.) We showed in [**Mat2**] that $\rho_c((m, \tau), (m, \tau)) \geq 0$, and ρ_c satisfies the triangle inequality. Moreover, h_η and ρ_c are obviously Lipschitz. Since ρ_c is obviously symmetric, it follows that ρ_c is a pseudo-metric on $A_{L,c} = A_c = \{(m, \tau) \in M \times \mathbb{T} : \rho_c((m, \tau), (m, \tau)) = 0\}$.

We will call $A_{L,c}$ the *Aubry set associated to L and c*. In [**Mat2**], we showed that this is a non-empty compact sub-set of $M \times \mathbb{T}$. In [**Mat2**] we showed that if M is provided with a Riemannian metric and ρ_M is the associated metric, then there is a constant C such that

$$\rho_c((m_0, \tau_0), (m_1, \tau_1)) \leq C[\rho_M(m_0, m_1) + \|\tau_1 - \tau_0\|]^2, \qquad (\star)$$

for $(m_i, \tau_i) \in A_c, i = 0, 1$, where $\|\tau_1 - \tau_0\| = \inf \{|t_1 - t_0| : t_i \in \mathbb{R}, t_i \equiv \tau_i \pmod 1 (i = 0, 1)\}$.

To define the other sets in the diagram above, we choose a closed 1-form η on $M \times \mathbb{T}$ such that $[\eta]_M = c$ and $[\eta]_{\mathbb{T}} = -\alpha(c)$ (so that η is Mañé critical). We will say that an invariant probability measure μ on the phase space is *c-minimal* if it minimizes $\int (L - \hat{\eta})\, d\mu$ over the set of invariant probability measures on phase space. In view of the assumption that η is Mañé critical, this means that $\int (L - \hat{\eta})\, d\mu = 0$. Because $\int (L - \hat{\eta})\, d\mu = A_L(\mu) - \langle \rho(\mu), c \rangle + \alpha(c)$, the property of being c-minimal depends only on c and not on the choice of Mañé critical closed 1-form η such that $[\eta]_M = c$.

We let $\dot{M}_c = \dot{M}_{L,c}$ be the closure of the union of the supports of all c-minimal measures on $TM \times \mathbb{R}$. The *graph theorem* ([**Mat1**, Theorem 2]) asserts that $\pi | \dot{M}_c$ is an injective mapping of $\dot{M}_c$ into $M \times \mathbb{T}$ and that its inverse $\pi^{-1} : \pi(M_c) \longrightarrow \dot{M}_c$ is Lipschitz. We set $M_c = M_{L,c} = \pi(\dot{M}_c)$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

We say that an absolutely continuous curve $\gamma : \mathbb{R} \longrightarrow M$ is a *c-minimizer* if, for any interval $[a, b]$ and any absolutely continuous curve $\gamma_1 : [c, d] \longrightarrow M$ such that $c - a \in \mathbb{Z}$, $d - b \in \mathbb{Z}$, $\gamma_1(c) = \gamma(a)$, and $\gamma_1(d) = \gamma(b)$, we have

$$\int_a^b (L - \hat{\eta})(\gamma(t), \dot{\gamma}(t), t)\, dt \leq \int_c^d (L - \hat{\eta})(\gamma_1(t), \dot{\gamma}_1(t), t)\, dt,$$

where η is a Mañé critical closed 1-form on $M \times \mathbb{T}$ such that $[\eta]_M = c$. This is the same as the notion introduced in [**Mat2**, §5]. (Note, however, that throughout [**Mat2**], $\alpha(c)$ has the wrong sign in front of it.) We define $\dot{N}_{L,c} = \dot{N}_c = \bigcup\{(\gamma(t), \dot{\gamma}(t), t \mod 1) : \gamma \text{ is a } c\text{-minimizer and } t \in \mathbb{R}\}$. Thus $\dot{N}_c \subset TM \times \mathbb{T}$.

Let $\gamma : \mathbb{R} \longrightarrow M$ be a c-minimizer. Let α, α' be in the α-limit set of γ and let ω, ω' be in the ω-limit set of γ. From the discussion in [**Mat2**, §6], it follows that $\rho_c(\alpha, \alpha') = \rho_c(\omega, \omega') = 0$ and $\rho_c(\alpha, \omega) = \rho_c(\alpha', \omega')$. If $\rho_c(\alpha, \omega) = 0$, we say that γ is a *regular* c-minimizer. We let $\dot{A}_c = \bigcup\{(\gamma(t), \dot{\gamma}(t), t \mod 1) : \gamma \text{ is a regular } c\text{-minimizer and } t \in \mathbb{R}\}$.

By definition $\dot{A}_c \subset \dot{N}_c$. It follows from [**Mat2**, Theorem 6.1] that $\dot{M}_c \subset \dot{A}_c$, $\pi(\dot{A}_c) = A_c$ and the graph theorem has the following extension: the mapping $\pi : \dot{A}_c \longrightarrow A_c$ is a bi-Lipschitz homeomorphism. (Note, however, that in the penultimate sentence before the statement of Theorem 6.1, the word 'regular' was inadvertently omitted. This sentence should read: 'We let $M'_c \subset TM \times \mathbb{T}$ denote the union of all regular c-minimal trajectories.')

This completes the definition of the sets that appear in the diagram above. In addition, we define $(\bar{A}_{L,c}, \rho_{L,c})$ to be the metric space obtained by identifying two points in the Aubry set $A_{L,c}$ if their $\rho_{L,c}$-pseudo-distance vanishes. We call $(\bar{A}_{L,c}, \rho_{L,c})$ the *quotient Aubry metric space* and $\bar{A}_{L,c}$ the *quotient Aubry set*. Clearly the quotient Aubry metric space is compact.

In the autonomous case, one may make the same definitions, of course. In this case, the $\mathbb{T}$-factor is irrelevant and we may summarize the discussion of action minimizing sets by the following commutative diagram:

$$\begin{array}{ccccccc}
\dot{M}_{L,c} & \subset & \dot{A}_{L,c} & \subset & \dot{N}_{L,c} & \subset & TM \\
\downarrow & & \downarrow & & & & \downarrow \pi \\
M_{L,c} & \subset & A_{L,c} & & \subset & & M
\end{array}$$

The only difference between this and the previous commutative diagram is that the $\mathbb{T}$-factor is left out from the right-most column. An autonomous system is time-periodic; for such a system the sets in the previous diagram may be obtained from the sets above by multiplying with $\mathbb{T}$.

In the autonomous case, the functions $h_\eta((m_0, \tau_0), (m_1, \tau_1))$ and $\rho_c((m_0, \tau_0), (m_1, \tau_1))$ are independent of τ_0 and τ_1, so we may write $h_\eta(m_0, m_1) = h_\eta((m_0, \tau_0), (m_1, \tau_1))$ and $\rho_c(m_0, m_1) = \rho_c((m_0, \tau_0), (m_1, \tau_1))$. Fathi's result [**F**] on the convergence of the Lax–Oleinik semi-group in the autonomous case implies that

$$h_\eta(m_0, m_1) = \lim_{T \to +\infty} \left\{ \inf_\gamma \int_0^T (L - \widehat{\eta})(\gamma(t), \dot{\gamma}(t))\, dt \right\},$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

where γ ranges over all absolutely continuous curves $\gamma : [0, T] \to M$ such that $\gamma(0) = m_0$ and $\gamma(T) = m_1$. The non-trivial point here is that the limit exists. Note that the corresponding limit in the time-periodic case does not exist [**FaM**]. It also follows from Fathi's result that in the autonomous case it is possible to remove the conditions $c - a \in \mathbb{Z}$ and $d - b \in \mathbb{Z}$ from the definition of c-minimizer without thereby changing the class of trajectories that are c-minimizers.

The examples that we construct in this paper are all autonomous. However, we have reviewed the time-periodic case in this section because we wish to point out what our examples mean in this setting.

2. *The main result*

The main result of this paper is the construction of certain examples of Lagrangians. In these examples, the state space M is a d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Let $\theta_1, \dots, \theta_d$ (mod 1) denote the standard coordinates on $\mathbb{T}^d$. Let $\theta_1, \dots, \theta_d, \dot\theta_1, \dots, \dot\theta_d$ denote the associated system of coordinates on the tangent space of $\mathbb{T}^d$. Our first examples have the form

$$L_0 = \tfrac{1}{2} \sum \dot\theta_i^2 + U(\theta_1, \dots, \theta_d) \qquad (\star)$$

where U is a non-negative function on $\mathbb{T}^d$ that vanishes somewhere. (In particular, these Lagrangians are autonomous.) We let Z_U denote the zero set of U. It is easily seen that $A_{L_0,0} = Z_U$, where 0 denotes the zero cohomology class in $H^1(\mathbb{T}^d; \mathbb{R})$.

We say that a function is of class $C^{r,\alpha}$ if it is r-times continuously differentiable and its rth derivative satisfies a Hölder condition of order α. Here is the main result.

THEOREM 1. *Let $\epsilon > 0$, $d \geq 2$. It is possible to choose U of class $C^{2d-3,1-\epsilon}$ such that $(\bar A_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval.*

Theorem 1 provides examples in every dimension $d \geq 2$. However, in the case $d = 2$, the Lagrangian L_0 is only of class $C^{1,1-\epsilon}$, so it does not formally fit the definition of 'Lagrangian' that we used in previous papers, starting with [**Mat1**], since we required a 'Lagrangian' to be C^2. In our previous papers this requirement was intended only as a convenience to permit us to formulate the Legendre condition. As it happens, however, our main result in [**Mat4**] depends crucially on the hypothesis that the Lagrangian is C^2. There we showed that if L is a (C^2) Lagrangian (even time-dependent) whose state space has dimension less than or equal to 2, then the quotient Aubry set $\bar A_{L,c}$ associated to a cohomology class $c \in H^1(M; \mathbb{R})$ is totally disconnected. Of course, the case $d = 2$ of Theorem 1 shows that the assumption that the differentiability class of L is C^2 cannot be weakened.

In §5, we will show that Theorem 1 is a consequence of the Proposition in §5. We will finish the proof of this Proposition in §14.

It is not possible to improve the differentiability of U in Theorem 1. In other words, if U is of class C^{2d-2}, then $\bar A_{L_0,0}$ is totally disconnected, as we show in §19. However, this result depends on L_0 having the form ($\star$). For general (time-dependent or autonomous) Lagrangians, we know of no results concerning the total disconnectedness of $\bar A_{L,c}$ when

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

the state space has dimension less than 2. In particular, we do not know the solution to the following.

Problem. If a Lagrangian L is C^∞, is $\bar{A}_{L,c}$ totally disconnected for every $c \in H^1(M;\mathbb{R})$ (where M denotes the state space of L)?

Our constructions permit the construction of Riemannian manifolds with interesting quotient Aubry sets.

THEOREM 2. *Let $\epsilon > 0, d \geq 3$. It is possible to choose a Riemannian metric g of class $C^{2d-5,1-\epsilon}$ on a d-dimensional torus $\mathbb{T}^d$ and a cohomology class $c \in H^1(\mathbb{T}^d;\mathbb{R})$ such that $(\bar{A}_{K,c}, \rho_{K,c})$ is isometric to an interval, where K is the kinetic energy associated to g.*

Note that c is necessarily non-zero, since $A_{K,0} = \mathbb{T}^d$ so $\rho_{K,0} \equiv 0$ is in view of ($\star$) in §1.

In the case $d = 3$, the metric so constructed is of class $C^{1,1-\epsilon}$. Proposition 2 of [**Mat4**] shows that it is impossible to have such an example of class C^2. (The statement of this Proposition does not explicitly contain a differentiability hypothesis, but we did assume at the outset of [**Mat4**] that all Lagrangians were C^2. Since this is a standing assumption throughout [**Mat4**], it means that the Riemannian metric mentioned in Proposition 2 is of class C^2.)

In the case $d > 3$, we do not know whether it is possible to improve Theorem 2 by finding examples of higher differentiability class.

In §15, we will use the Maupertuis–Carneiro principle to deduce Theorem 2 from Theorem 1.

Theorem 1 also provides a partial converse of the remark at the end of [**Mat4**]. There we observed that if L is an autonomous Lagrangian (of class C^2) whose state space M has dimension less than or equal to 3 and $c \in H^1(M;\mathbb{R})$ is such that $\dot{A}_c$ does not meet the zero section of TM, then $\bar{A}_{L,c}$ is totally disconnected. On the other hand, Theorem 1 provides an example of an autonomous Lagrangian L of class $C^{3,1-\epsilon}$ whose state space is $\mathbb{T}^3$ such that $\bar{A}_{L,c}$ is isometric to an interval. In this case, $\dot{A}_c$ is in the zero section of TM.

Let L_0 be given by ($\star$) and set

$$L_1 = L_0 + V(\theta) \cdot \dot{\theta},$$

where $V = (V_1, \ldots, V_d)$ is a vector-valued function of $\theta = (\theta_1, \ldots, \theta_d)$. In §18, we will show that, for a suitable choice of V (of class C^∞), the set-valued function $L \mapsto A_{L,0}$ is not upper semi-continuous at L_1 on the space of Lagrangian L of class $C^{2d-3,1-\epsilon}$ provided with the $C^{2d-3,1-\epsilon}$ topology.

We will also prove a converse of [**Mat3**, Proposition 1].

THEOREM 3. *Let $d \geq 2$. There exists a compact, connected sub-set Z of $\mathbb{R}^d$ and $x, y \in Z$ with the following property. Let $\beta < d$. There exists $\delta > 0$ such that for any sequence $x = x_0, \ldots, x_n = y$ of points in Z, we have*

$$\sum_{i=1}^{n} \|x_i - x_{i-1}\|^\beta \geq \delta.$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Theorem 3 justifies the statement following [**Mat3**, Proposition 1], 'The exponent d cannot be improved'. I recently learned that there is an example in an earlier paper [**A**] that also justifies this statement. I am grateful to Albert Fathi for pointing out Assouad's result to me.

We will prove Theorem 3 in §4.

3. *Construction of Z*

The main part of this paper is the construction of a function U that has the properties stated in Theorem 1. Our first step towards constructing U is to construct the zero set $Z = Z_U$ of U. We carry out this construction in this section.

We begin with some notation and terminology. If A and B are two sub-sets of $\mathbb{R}^d$, we set $A + B = \{a + b : a \in A,\ b \in B\}$. If $\lambda \in \mathbb{R}$, we set $\lambda A = \{\lambda a : a \in A\}$. By the *standard cube* Q we mean

$$Q = \{(x_1, \ldots, x_d) \in \mathbb{R}^d : 0 \leq x_i \leq 1, \text{ for } i = 1, \ldots, d\}.$$

By an nth-*order dyadic cube* in $\mathbb{R}^d$ we mean a set of the form

$$2^{-n}(a + Q), \quad a \in \mathbb{Z}^d.$$

By the nth-*order dyadic grid* in $\mathbb{R}^d$ we mean the family of all nth-order dyadic cubes. This family covers $\mathbb{R}^d$ and the intersection of two distinct members of it is contained in the boundary of each. By an nth-*order dyadic vertex, edge*, or k-*face*, we mean a vertex, edge, or k-dimensional face of an nth-order dyadic cube.

We will use a similar terminology concerning sub-sets of the d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Thus, an nth-*order dyadic cube* in $\mathbb{T}^d$ is the image of an nth-order dyadic cube in $\mathbb{R}^d$ under the projection $\mathbb{R}^d \longrightarrow \mathbb{T}^d$. By the nth-*order dyadic grid* in $\mathbb{T}^d$ we mean the family of all nth-order dyadic cubes in $\mathbb{T}^d$, and so on.

By an nth-*order dyadic line segment* in $\mathbb{R}^d$ or $\mathbb{T}^d$ we mean a (straight) line segment that is a union of nth-order dyadic edges. By an nth-*order dyadic rectangle* we mean a rectangle that is a union of nth-order dyadic 2-faces. We will speak of the *top*, *bottom*, and two (*vertical*) *sides* of such a rectangle. The meaning of these terms needs no explanation; we only remark that our choice of which axis of the rectangle is horizontal and which is vertical depends on the context and will be specified in each instance. Likewise, the direction of the vertical (needed to distinguish the top from the bottom) depends on the context and will also be specified. Usually, we will make these specifications by specifying the bottom of the rectangle.

By the *length* of a line segment we mean its Euclidean length. By the nth-*order edge-length* of an nth-order dyadic line segment we mean the number of nth-order dyadic edges that it contains.

Let J be an mth-order dyadic line segment and let $n \geq m$. An (m, n)th-*order dyadic comb* C *with base* J is constructed as follows.

The comb has a number of *teeth*. Each tooth of the comb is an nth-order dyadic rectangle whose bottom is an mth-order dyadic edge of J. Taking the mth-order dyadic edges of J in order beginning at one end of J, every other edge (starting with

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

the first one) is the bottom of a tooth, with the exception of the next-to-last edge, which is not the bottom of a tooth. The comb C is J with the bottoms of the teeth removed, and the sides and top of the teeth added. All the teeth point in the same direction, i.e. they all lie in the same plane and, relative to that plane, they all lie on the same side of J. All sides of all teeth have the same length.

Thus, C is a Jordan curve and a union of nth-order dyadic edges. The odd-numbered mth-order dyadic edges of J, with the exception of the next-to-last edge, are the bottoms of teeth and are not part of C. The even-numbered mth-order dyadic edges and the next-to-last edge are not the bottoms of teeth and are part of C.

The base J of C is parallel to a coordinate axis of $\mathbb{R}^d$. We call this the *primary axis* of C. Every side of a tooth of C is parallel to a second coordinate axis and all of these are parallel to the same axis. We call this second axis the *secondary axis* of C. Each tooth of C has a top and a bottom. We call the direction from the bottom to the top the *positive orientation* of the secondary axis. Since all teeth lie on the same side of J (relative to the plane in which they lie), it follows that the direction is independent of the tooth.

We call the length of J the *primary length* of C and the length of a side of a tooth the *secondary* (or *ultimate*) *length* of C.

To construct Z, we will iterate the construction of a comb infinitely often and take the limit. Now we describe the first iteration.

Let C be an (m, n)th-order dyadic comb with base J. Let $p \geq n$. Let $J_1, \ldots, J_{2K}$ be the sides of the teeth of C. Each J_i is an nth-order dyadic line segment. We replace each J_i by an (n, p)th-order dyadic comb C_i with base J_i.

For this purpose, we take the bottom vertex of J_i to be the first vertex. (This means that the odd-numbered nth-order dyadic edges of J_i beginning at the bottom of J_i, with the exception of the next-to-last edge, are the bottoms of teeth of C_i. In particular, the last three nth-order dyadic edges of J_i before the top of J_i are not the bottoms of teeth of C_i.) We choose the C_i's to all have the same secondary length. (They all have the same primary length, because the primary length of C_i is the length of J_i.)

We consider two cases, depending on whether $d = 2$ or $d > 2$.

In the case $d > 2$, we take all the C_i to have the same secondary axis, different from the primary axis of C, and we require that all C_i lie on the same side of the affine plane that contains C, relative to the affine 3-space that contains C and all the C_i's.

In the case $d = 2$, there are just two coordinate axes. One is the primary axis of C; the other is the secondary axis of C. Since the secondary axis of C is the primary axis of C_i, it follows that the secondary axis of C_i must be the primary axis of C. Let R_j be the tooth of C one of whose sides is the base J_i of C_i. We require that $C_i \subset R_j$. In addition, we require that the length of a side of a tooth of C_i be less than or equal to 2^{-m-6}.

The significance of the condition $C_i \subset R_j$ is that it tells which side of J_i the comb C_i lies on.

Each tooth R_i of C has two sides; call them J_k and J_j. The distance between J_k and J_j is 2^{-m} since they are separated by an mth-order dyadic edge. We have $C_k \subset R_i$ and $C_j \subset R_i$; the condition that the length of a tooth of C_k (or C_j) is less than or equal to 2^{-m-6} implies that the distance between C_k and C_j is greater than or equal to $2^{-m} - 2 \times 2^{-m-6}$ which is greater than or equal to 2^{-m-1}.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

We let C' be C with each J_i replaced by C_i. We call C' an (n, p)th-*order dyadic once iterated comb*. It is a Jordan arc and a union of pth-order dyadic edges.

We call J the *primary base* and C the *secondary* (or *ultimate*) *base* of C'. By the *primary* and *secondary lengths* of C', we mean the primary and secondary lengths of C; by the *ternary* (or *ultimate*) *length* of C', we mean the secondary length of any one of the C_i's; by the definition given above, they are all the same. Note also that, by the definition given above, the secondary length of C' (or of C) is the same as the primary length of any one of the C_i's.

Continuing in this way, we construct twice iterated, thrice iterated, ..., m-times iterated, $\cdots$ combs. Thus, we obtain sequences $J, C, C', C'', \ldots, C^{(m)}, \ldots$ where J is a compact line segment, C is a comb with base J, C' is a once iterated comb with ultimate base C', C'' is a twice iterated comb with ultimate base C', and so on. Next, we describe the inductive step in this construction.

For the inductive step, we suppose that a sequence $J, C, C', \ldots, C^{(m)}$ of sub-sets of $\mathbb{R}^d$ has been constructed, where J is an n_0th-order dyadic line segment, C is a (n_0, n_1)th-order dyadic comb with base J, C' is an (n_1, n_2)th-order dyadic once iterated comb with ultimate base C, and so on. The aim of our inductive definition is to define what we mean by an (n_{m+1}, n_{m+2})th-*order dyadic* $(m+1)$*-times iterated comb* $C^{(m+1)}$ *with ultimate base* $C^{(m)}$. For this purpose, we make the inductive assumption that $C^{(k)}$ is an (n_k, n_{k+1})th-order dyadic k-times iterated comb with ultimate base $C^{(k-1)}$ for $0 \leq k \leq m$. In the case $k = 0$, this means that $C^{(0)} = C$ is an (n_0, n_1)th-order dyadic comb with base $C^{(-1)} = J$; in the case $k = 1$, it means that $C^{(1)} = C'$ is an (n_1, n_2)th-order dyadic comb with ultimate base $C^{(0)} = C$; and, in general, its meaning is given by the inductive step below. We will also use the notion of the *ultimate teeth* of $C^{(k)}$. In the case $k = 0$, this means the teeth of C; in the case $k = 1$, it means the secondary teeth of C'; and in general, its meaning is again given by the inductive step below.

Here is the inductive step in the definition of m-times iterated comb:

We suppose $n_{m+2} > n_{m+1} + 1$ is an integer. We let $J_1, \ldots, J_{2k}$ denote the sides of the ultimate teeth of $C^{(m)}$. We define $C^{(m+1)}$ to be $C^{(m)}$ modified by replacing each J_i by an (n_{m+1}, n_{m+2})th-order dyadic comb C_i whose base is J_i. By induction, $C^{(m)}$ is an (n_m, n_{m+1})th-order dyadic m-times iterated comb, so each J_i is an n_{m+1}th-order dyadic line segment. We take the bottom vertex of J_i to be the first vertex. This has the same consequence as before. The last three n_{m+1}th-order dyadic edges of J_i before the top of J_i are not bottoms of teeth of C_i. This fact simplifies certain estimates that we need later.

From the definition (just given) of $C^{(m+1)}$, it follows that

$$C^{(m+1)} = [C^{(m)} \setminus (J_1 \cup \cdots \cup J_{2k})] \cup (C_1 \cup \cdots \cup C_{2k}).$$

We will call such a $C^{(m+1)}$ an $(m+1)$*-times iterated comb*, provided that $C^{(m)}$ is an m-times iterated comb and the conditions on the combs $C_1, \ldots, C_{2k}$ stated below are satisfied. Note that, in the cases $m = 0$ and $m = 1$, this agrees with our previous definitions of comb and once-iterated comb. We call the teeth of the combs $C_1, \ldots, C_{2k}$ the *ultimate teeth* of $C^{(m+1)}$.

The first conditions that we require the C_i's to have is that they all have the same secondary lengths and they all have the same secondary axis. We call these the *ultimate*

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

length and *ultimate axis* of $C^{(m+1)}$. Note that all the C_i's have the same primary length and same primary axis by induction: the primary length of C_i is the length of J_i, which is the ultimate length of $C^{(m)}$; and the primary axis of C_i is the axis parallel to J_i, which is the ultimate axis of $C^{(m)}$.

We call the length and axis of J the *primary* length and axis of $C^{(m+1)}$, the secondary (or ultimate) length and axis of C the *secondary* length and axis of $C^{(m+1)}$, the ternary (or ultimate) length and axis of C' the *ternary* length and axis of $C^{(m+1)}$, and generally the ultimate length and axis of $C^{(k)}$ the $(k+2)$-*ary* length and axis of $C^{(m+1)}$ (for $k \leq m+1$).

In the case $m+2 < d$, we require that all the axes of $C^{(m+1)}$ (i.e. the primary axis, secondary axis, ..., ultimate axis) be different.

There are $(m+3)$ such axes. It follows that $C^{(m+1)}$ lies in a unique affine $(m+3)$-dimensional sub-space $S^{(m+1)}$ of $\mathbb{R}^d$. We require that all of the ultimate teeth of $C^{(m+1)}$ lie on the same side of $S^{(m)}$ relative to $S^{(m+1)}$.

In the case that $m+2 = d$, every axis of $\mathbb{R}^d$ is already an axis of $C^{(m)}$, so the ultimate axis of $C^{(m+1)}$ must be one of the axes of $C^{(m)}$. We require that it be the primary axis.

Let $S(= S^{(0)})$ be the affine plane in $\mathbb{R}^d$ that contains C. Let $\pi : \mathbb{R}^d \longrightarrow S$ be the orthogonal projection. Let $R_1, \ldots, R_\ell$ be the teeth of C, with sides $J_1, \ldots, J_{2\ell}$. Let R^* be an ultimate tooth of $C^{(m+1)}$. We require that $\pi(R^*)$ lie in one of the R_i's (in the case that $m+2 = d$). Note that in the case $d = 2$, this requirement agrees with a requirement in our construction above of C'. In the case $d > 2$, it follows by induction from the requirements that we have already imposed that $\pi(R^*)$ is a line segment one of whose ends (the projection of the bottom of R^*) is in one of the J_j's. Thus, the requirement that $\pi(R^*)$ lie in one of the R_i's specifies on which side of $\pi^{-1}(J_j)$ the tooth R^* lies.

Finally (still in the case $m+2 = d$), we require that the ultimate length of $C^{(m+1)}$ (i.e. the length of a side of a tooth of one of the C_i's) be less than or equal to 2^{-n_0-6}. (This generalizes a condition we imposed above in the case $d = 2$.)

In order to state the further condition on $C_1, \ldots, C_{2k}$ that we require in the case $m+2 > d$, we introduce the following notion.

As before, we let $J_1, \ldots, J_{2\ell}$ be the sides of the teeth of C. For $-1 \leq m' \leq m$ and $1 \leq j \leq 2\ell$, we define $C_j^{(m')}$ by induction on m' as follows. First, $C_j^{(-1)} = J_j$. Second,

$$C_j^{(m')} = [C_j^{(m'-1)} \smallsetminus (J_1' \cup \cdots \cup J_{2\alpha}')] \cup (C_1' \cup \cdots \cup C_{2\alpha}'),$$

where J_i' runs over the sides of the ultimate teeth of $C^{(m')}$ with $J_i \subset C_j^{(m'-1)}$ and C_i' is the comb that replaces J_i' in the construction of $C^{(m'+1)}$.

Thus, $C_j^{(m')} \subset C^{(m'+1)}$. It is easy to see that, for $m+2 \leq d$, we have that $C_j^{(m')}$ is an m'-times iterated comb with ultimate base $C_j^{(m'-1)}$, if $0 \leq m' \leq m$.

We require that this also hold in the case $m+2 > d$. It is easy to see that this amounts to a further condition on $C_1, \ldots, C_{2k}$. In fact, by inductive hypothesis, it holds for $0 \leq m' < m$, so this is an additional condition (for the inductive definition of what it means for $C^{(m+1)}$ to be an $(m+1)$-times iterated comb) only in the case $m' = m$. In the case that $m' = m$ and $m+2 > d$ the statement that $C_j^{(m)}$ is an m-times iterated comb amounts to a restriction of the placement of those C_i that are in $C_j^{(m)}$. This restriction is provided by the previous

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

construction when $m+2=d+1$ and by induction when $m+2>d+1$. Note that this provides a restriction on the placement of every C_i, since every C_i is in some $C_j^{(m)}$.

This completes our inductive definition of m-times iterated comb. We call J the *primary base* of $C^{(m+1)}$, C the *secondary base*, C' the *ternary base*, ..., $C^{(k)}$ the $(k+2)$-*ary base*, ..., and $C^{(m)}$ the $(m+2)$-ary or *ultimate base*.

All such iterated combs are Jordan curves. This is immediate from the definition when $m+2<d$.

When $m+2=d$, we have $\pi^{-1}(J_j)\cap C^{(m)}=C_j^{(m-1)}$, where (as above) $\pi:\mathbb{R}^d\to S$ is the orthogonal projection. Moreover, if R_i is a tooth of C with sides J_j and J_k, then

$$\pi^{-1}(R_i)\cap C^{(m+1)}=C_j^{(m)}\cup C_k^{(m)}\cup T_i,$$

where T_i is the top of R_i.

The last requirement that we imposed in the case $m+2=d$ implies that the distance between $C_j^{(m)}$ and $C_k^{(m)}$ $\geq 2^{-n_0}-2\times 2^{-n_0-6}>2^{-n_0-1}$. Thus $C_j^{(m)}$ and $C_k^{(m)}$ do not overlap. It follows that $C^{(m+1)}$ is a Jordan curve.

When $m+2>d$, it follows by induction that $C^{(m+1)}$ is a Jordan curve.

Now, we consider an infinite sequence $J, C, C', C'', \ldots, C^{(m)}, \ldots$ where J is an n_0th-order dyadic line segment, C is an (n_0,n_1)th-order dyadic comb with base J, and, for $m\geq 1$, $C^{(m)}$ is an (n_m,n_{m+1})th-order dyadic m-times iterated comb with ultimate base $C^{(m-1)}$. We let ℓ_0 denote the length of J, ℓ_1 the secondary (or ultimate) length of C, and generally ℓ_m the ultimate length of $C^{(m-1)}$. From our construction,

$$\ell_{m+d}\leq 2^{-n_m-6},\quad m\geq 0.$$

In the case $m=0$, this follows from our construction of $C^{(d-1)}$. In the case that $m\geq 1$, it follows by induction on m.

Conversely, if $2\leq n_0<n_1<n_2<\cdots$ are positive integers with $n_m+1<n_{m+1}$ and ℓ_m is the product of 2^{-n_m} by an even integer greater than or equal to 4 for $m\geq 0$, then it is easy to see that there is a sequence $J, C, C', \ldots, C^{(m)}, \ldots$, as above, i.e. $C^{(m)}$ is an (n_m,n_{m+1})th-order m-times iterated comb whose ultimate length is ℓ_m.

It is easy to see that

$$J\cap C\subset C\cap C'\subset\cdots\subset C^{(m-1)}\cap C^{(m)}\subset\cdots.$$

We set

$$C^{(\infty)}=\bigcup_m C^{(m-1)}\cap C^{(m)}.$$

We have

$$C^{(\infty)}=\lim_{m\to\infty}C^{(m)},$$

where the limit is taken in the Hausdorff topology. It is a Jordan curve.

We define Z to be such a $C^{(\infty)}$ for the special case where

$$n_m=[(m+10)^2/2d]+10+\sigma(m+1),\quad m\geq 0,$$
$$\ell_m=2^{-n_m+5+m}=2^{-[(m+10)^2/2d]-5-\sigma-(\sigma-1)m},\quad m\geq 0,$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

where $[x]$ denotes the integral part of x, i.e. the greatest integer less than or equal to x, and $\sigma = \sigma(d)$ is a sufficiently large integer, depending only on d. (It will be enough to suppose that $\sigma(d) \geq \max(5, d)$.) Note that $n_{m+1} \geq n_m + \sigma$ for all m.

Naturally, the exact numbers here are not important. What is important is that n_m grows asymptotically like $m^2/2d$ and the ultimate n_mth-order edge length of $C^{(m-1)}$ (i.e. the number of n_mth-order dyadic edges contained in a side of an ultimate tooth of $C^{(m-1)}$) grows proportionately to 2^m. We will also need later that the projection of Z into $\mathbb{T}^d$ is an isometry. For this, $\ell_m \leq 1/2$, for all m, is enough. In fact, we have $\ell_m \leq 2^{-10}$ for the choice we have made. Finally, we need $\ell_{m+d} \leq 2^{-n_m-6}$, for $m \geq 0$.

To verify this, we observe that $n_{m+d} = [(m+10)^2/2d + m + 10 + d/2] + 10 + \sigma(m + d + 1) = n_m + 10 + m + (2\sigma + 1)d/2 + \Delta$, where $\Delta = 0$ if d is even and $\Delta = \pm 1/2$ if d is odd. Hence $\ell_{m+d} = 2^{-n_m-5-(2\sigma+1)d/2-\Delta}$. Since $d/2 + \Delta \geq 1$ if $d \geq 2$, we have $\ell_{m+d} \leq 2^{-n_m-6}$, as asserted. We also have $\ell_{m+d} \leq 2^{-n_m-\sigma-3}$.

4. *A property of Z*

Let $\beta \geq 1$ and let $X \subset \mathbb{R}^d$. For $x, y \in \mathbb{R}^d$, we set $\sigma(x, y) = \sigma_{\beta,X}(x, y) = \inf\{\Sigma_{i=1}^p \|x_i - x_{i-1}\|^\beta : x_i \in X \text{ for } 0 \leq i \leq p, x_0 = x, \text{ and } x_p = y\}$.

If $\beta \geq d$ and X is compact and connected, then $\sigma_{\beta,X}(x, y) = 0$, whatever the choice of x and y in X. For $\beta > d$, this is easy to see. For $\beta = d$, we proved this in [**Mat3**].

In this section, we will show that this is false for $\beta < d$. We will show that the set Z that we constructed in the last section is a counter-example. Other counter examples are given in [**A**]. Thus, we will prove the following.

PROPOSITION. *There exist $x, y \in Z$ such that if $\beta < d$, then $\sigma_{\beta,Z}(x, y) > 0$.*

Theorem 3 in §2 is an immediate consequence of this Proposition.

The rest of this section is a proof of this Proposition.

Let $w_m = 2^{-n_m}$ denote the length of an n_mth-order dyadic edge and let ν_m denote the number of n_{m+1}th-order dyadic edges in a side of an ultimate tooth of $C^{(m)}$ (i.e. the n_{m+1}th-order edge length of such a side). Since the length of such a side of an ultimate tooth is $\ell_{m+1} = 2^{-n_{m+1}+6+m}$, we have that $\nu_m = 2^{6+m}$.

Our first step towards showing that $\sigma_{\beta,Z}$ is not identically zero when $\beta < d$ is to show that there exists a positive integer $m_0 = m_0(\beta)$ such that

$$(\nu_m/4)(w_{m+1}/2)^\beta \geq (w_m/2)^\beta \quad \text{for } m \geq m_0.$$

Since $\nu_m = 2^{6+m}$ and $w_m = 2^{-n_m} = 2^{-[(m+10)^2/2d]-10-\sigma(m+1)}$, this inequality is equivalent to

$$2^{4+m-\beta([(m+11)^2/2d]+11+\sigma(m+2))} \geq 2^{-\beta([(m+10)^2/2d]+11+\sigma(m+1))},$$

or

$$4 + m - \beta([(m+11)^2/2d] + 11 + \sigma(m+2)) \geq -\beta([(m+10)^2/2d] + 11 + \sigma(m+1)).$$

It is enough to verify that

$$4 + m \geq \beta\left(\frac{m}{d} + \frac{21}{2d} + \sigma + 1\right).$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Since $\beta < d$, this clearly holds if m is large enough. This proves the inequality above, which in turn implies

$$(\nu_m^*/2)(w_{m+1}/2)^\beta \geq (w_m/2)^\beta \quad \text{for } m \geq m_0, \tag{$\star$}$$

where $\nu_m^* = \nu_m - 2$.

Let R be an ultimate tooth of $C^{(m)}$ and let x and y be the endpoints of the bottom of R. In what follows, we will assume that $m \geq m_0$. Under this assumption, we will prove that $\sigma_{\beta,Z}(x, y) \geq (w_m/2)^\beta$.

To prove this, it is enough to consider a sequence $x = x_0, x_1, \ldots, x_p = y$ of points of Z and show that $\sum_{i=1}^p \|x_i - x_{i-1}\|^\beta \geq (w_m/2)^\beta$.

Let J_1 be the side of R containing x and J_2 the side of R containing y. Let T be the top of R. Then $J_1 \cup T \cup J_2$ is a Jordan sub-arc of $C^{(m)}$ whose endpoints are x and y. In the construction of $C^{(m+1)}$, each side of an ultimate tooth of $C^{(m)}$ is replaced by a comb. In particular, J_1 and J_2 are replaced by combs C_1 and C_2. Then $C_1 \cup T \cup C_2$ is a Jordan sub-arc of $C^{(m+1)}$ whose endpoints are x and y. More explicitly, C_1 is a Jordan sub-arc of $C^{(m+1)}$ whose endpoints are x and $\xi = J_1 \cap T$ and C_2 is Jordan sub-arc of $C^{(m+1)}$ whose endpoints are $\eta = T \cap J_2$ and y. We have that x, ξ, η, y are successive points in the Jordan arc $C^{(p)}$, for every $p \geq m$. Let $C_1^{(p-m-1)}$ be the sub-arc of $C^{(p)}$ joining x and ξ and $C_2^{(p-m-1)}$ be the sub-arc of $C^{(p)}$ joining η and y. Note that T is the sub-arc joining ξ and η, and that $C_i^{(p-m-1)}$ is a $(p-m-1)$-times iterated comb with primary base J_i $(i = 1, 2)$. Similarly, we let $C_1^{(\infty)}$ be the sub-arc of $C^{(\infty)}$ joining x and ξ and $C_2^{(\infty)}$ be the sub-arc of $C^{(\infty)}$ joining η and y. Again, T is the sub-arc of $C^{(\infty)}$ joining ξ and η. Moreover, $C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$ is the sub-arc of $C^{(\infty)}$ joining x and y. Thus, $C^{(\infty)} = A \cup (C_1^{(\infty)} \cup T \cup C_2^{(\infty)}) \cup B$, where A is a sub-arc of $C^{(\infty)}$ having x as an endpoint (or is x alone), B is a sub-arc of $C^{(\infty)}$ having y as an endpoint (or is y alone), $A \cap (C_1^{(\infty)} \cup T \cup C_2^{(\infty)}) = x$, and $B \cap (C_1^{(\infty)} \cup T \cup C_2^{(\infty)}) = y$.

To prove that $\sum_1^p \|x_i - x_{i-1}\|^\beta \geq (w_m/2)^\beta$, we may assume without loss of generality that $x_i \in C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$, for $i = 0, \ldots, p$. Assume that this is not the case. Then $x_i \in A \smallsetminus x$ or $x_i \in B \smallsetminus y$ for some i. Suppose, for example, that the former holds. Let $x_a, \ldots, x_b$ be a maximal set of successive x_i's in $A \smallsetminus x$. Thus, $x_i \in A \smallsetminus x$ for $a \leq i \leq b$ and $x_{a-1} \notin A \smallsetminus x$ and $x_{b+1} \notin A \smallsetminus x$. Let E be the n_mth-order dyadic edge in A that has x as one of its endpoints. If $y \in A \smallsetminus E$ and $z \in C_1^{(\infty)} \cup T \cup C_2^{(\infty)} \cup B$, then $\|y - z\| \geq 2^{-n_m} = w_m$. Thus, if $x_a \in A \smallsetminus E$, then $\|x_{a-1} - x_a\|^\beta \geq w_m^\beta > (w_m/2)^\beta$ and the sum has the desired lower bound. Likewise, if $x_b \in A \smallsetminus E$, $\|x_b - x_{b+1}\|^\beta > (w_m/2)^\beta$. Thus, we may assume that $x_a, x_b \in E \smallsetminus x$. In this case, $\|x_{a-1} - x_a\| \geq \|x_{a-1} - x\|$ and $\|x_b - x_{b+1}\| \geq \|x - x_{b+1}\|$ so $\Sigma_1^p \|x_{i-1} - x_i\|^\beta \geq \Sigma_1^{a-1} \|x_{i-1} - x_i\|^\beta + \|x_{a-1} - x\|^\beta + \|x - x_{b+1}\|^\beta + \Sigma_{b+2}^p \|x_{i-1} - x_i\|^\beta$, so we may replace the original sequence with a new sequence having fewer elements outside of $C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$, for which the sum is smaller. A similar argument applies in the case that some $x_i \in B \smallsetminus y$. Continuing in this way, we may replace the original sequence with new sequences having fewer and fewer elements outside of $C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$. With each replacement, the sum decreases. Eventually, we arrive at a sequence with all elements in $C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$ with a smaller sum than the original sequence. Thus, to get a lower bound for the sum, it is enough to suppose that all $x_i \in C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

LEMMA 1. *If no* $x_a \in T$, *then* $\sum \|x_i - x_{i-1}\|^\beta \geq (w_m/2)^\beta$.

Proof. In this case, $x_i \in C_1^{(\infty)} \cup C_2^{(\infty)}$ for $i = 0, \ldots, p$ and the result is a consequence of the facts that

$$\text{dist.}(C_1^{(\infty)}, C_2^{(\infty)}) \geq w_m/2$$

and $x \in C_1^{(\infty)}$, $y \in C_2^{(\infty)}$. □

Thus, to obtain the desired lower bound, we may assume that some $x_a \in T$. Let a be the smallest index for which this holds. Then $x_i \in C_1^{(\infty)} \cup C_2^{(\infty)}$ for $i = 0, \ldots, a-1$. If $x_i \in C_2^{(\infty)}$ for at least one of these i's, then $\text{dist}(C_1^{(\infty)}, C_2^{(\infty)}) \geq w_m/2$ implies the desired lower bound, so we may assume without loss of generality that $x_i \in C_1^{(\infty)}$, for $0 \leq i < a$.

Let $E_1, \ldots, E_{v_m}$ be the n_{m+1}th-order dyadic edges that lie in J_1, arranged in order from the bottom to the top of J_1. Let $x = v_0, v_1, \ldots, v_{v_m} = \xi$ be the n_{m+1}th-order dyadic vertices that lie in J_1, also arranged from the bottom to the top of J_1. The endpoints of E_i are v_{i-1} and v_i. For $1 \leq i \leq v_m^*/2$, E_{2i-1} is the bottom of an ultimate tooth R_i of $C^{(m+1)}$ and $E_{2i} \subset C^{(q)}$ for all $q \geq m$. For $i = v_m^*/2$, we have, in addition, $E_{2i+1}, E_{2i+2} \subset C^{(q)}$. Let $C_{1i}^{(\infty)}$ be the sub-interval of the Jordan curve $C_1^{(\infty)}$ whose endpoints are v_{2i-2} and v_{2i-1}. Thus, the sequence of points $v_0, \ldots, v_{v_m}$ decomposes the $C_1^{(\infty)}$ into sub-segments $C_{11}^{(\infty)}, E_2, C_{12}^{(\infty)}, E_4, \ldots, C_{1,v_m^*/2}, E_{v_m^*}, E_{v_m-1}, E_{v_m}$ (in this order).

Next, we show that

$$\sum_{j=1}^{a} \|x_j - x_{j-1}\|^\beta \geq \sum_{i=1}^{v_m^*/2} \min\left((w_{m+1}/2)^\beta, \sum_{j=1}^{p_i} \|y_{ij} - y_{i,j-1}\|^\beta\right) \qquad (\dagger)$$

for suitably chosen sequences $y_{i0}, \ldots, y_{i,p_i}$ in $C_{1i}^{(\infty)}$, with $y_{i0} = v_{2i-2}$ and $y_{i,p_i} = v_{2i-1}$.

Our method for showing this is to use sub-sums of the left-hand side to dominate one or more summands on the right-hand side. First, we consider the possibility that $x_{j-1}, v_k, v_\ell, x_j$ occur in that order on the Jordan curve $C_1^{(\infty)}$, for some $k < \ell$. In this case, $\|x_j - x_{j-1}\| \geq \|v_\ell - v_k\| = (\ell - k)w_{m+1}$, and the single term $\|x_j - x_{j-1}\|^\beta$ on the left-hand side is greater than or equal to $(\ell - k + 1)(w_{m+1}/2)^\beta$, so it dominates the sum of the terms corresponding to indices i such that at least one endpoint of $C_{1i}^{(\infty)}$ is a v_m with $k < m < \ell$. Thus, we may eliminate these summands from further consideration. Likewise in the case that $x_j, v_k, v_\ell, x_{j-1}$ occur in that order on the Jordan curve $C_1^{(\infty)}$, we may eliminate the same summands.

Now we consider an index i for which the corresponding summand has not been eliminated in this way. In this case there exist $1 \leq j_0 < j_1 \leq p$ such that $x_{j_0} \in E_{2i-2}$ (or $x_{j_0} = v_0 = x$, in the case $i = 1$), $x_j \in C_{1i}^{(\infty)}$ for $j_0 < j < j_1$ and $x_{j_1} \in E_{2i}$. Let $y_{ij} = x_{j_0+j}$ for $1 \leq j \leq j_1 - j_0 - 1$, $y_{i0} = v_{2i-2}$, $y_{ip_i} = v_{2i-1}$, where $p_i = j_1 - j_0 < p$. Clearly,

$$\sum_{j=j_0+1}^{j_1} \|x_j - x_{j-1}\|^\beta \geq \sum_{j=1}^{p_i} \|y_{ij} - y_{i,j-1}\|^\beta,$$

so the ith-summand is dominated. It is easy to see that the different summands on the right-hand side are dominated by different sub-sums on the left, so (†) is proved.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

LEMMA 2. *Let $m \geq m_0$, let R be an ultimate tooth of $C^{(m)}$, let x and y be the endpoints of the bottom of R, and let $x = x_0, x_1, \ldots, x_p = y$ be a sequence of points of Z. Then*

$$\sum_{i=1}^{p} \|x_i - x_{i-1}\|^\beta \geq (w_m/2)^\beta.$$

Proof by induction on p. We showed above that it is no loss of generality to assume that every $x_i \in C_1^{(\infty)} \cup T \cup C_2^{(\infty)}$. According to Lemma 1, the conclusion holds if no $x_a \in T$. In particular, this is the case if $p = 1$, so the conclusion holds in this case. Thus, we may assume $p > 1$ and some $x_a \in T$. Let a be the smallest index for which this holds. We claim that the sums that appear under the main sum on the right-hand side of (†) satisfy

$$\sum_{j=1}^{p_i} \|y_{ij} - y_{i,j-1}\|^\beta \geq (w_{m+1}/2)^\beta,$$

by the induction hypothesis. This is because each sequence $y_{i0}, \ldots, y_{ip_i}$, satisfies the condition imposed on $x_0, x_1, \ldots, x_p$ in the hypothesis of Lemma 2, with respect to an ultimate tooth of $C^{(m+1)}$ (instead of $C^{(m)}$). Since $p_i < p$, we may apply the induction hypothesis.

In view of (†), we then obtain

$$\sum_{j=1}^{a} \|x_j - x_{j-1}\| \geq (v_m^*/2)(w_{m+1}/2)^\beta,$$

and (⋆) gives the desired conclusion. □

We now return to the proof of the Proposition. Let $m \geq m_0$, let R be an ultimate tooth of $C^{(m)}$, and let x and y be the endpoints of the bottom of R. Lemma 2 implies that $\sigma_{\beta,Z}(x, y) \geq (w_m/2)^\beta$. □

5. *Construction of U*

In this section, we construct a function U that has the properties stated in Theorem 1. However, the proof that it has those properties will not be finished until §14.

We let u be a C^∞ real-valued function on $\mathbb{R}$ that has the following properties:

- $0 \leq u \leq 1$;
- supp $u \subset (-\frac{1}{2}, \frac{3}{2})$, where supp u denotes the *support* of u (the closure of $\{u \neq 0\}$);
- u is identically 1 on $[0, 1]$;
- u is symmetric about $\frac{1}{2}$; and
- $u' \geq 0$ on $(-\frac{1}{2}, 0)$.

We let u_d be the real-valued function on $\mathbb{R}^d$ defined by

$$u_d(x_1, \ldots, x_d) = u(x_1) \cdots u(x_d).$$

Let C be a cube in $\mathbb{R}^d$ with center c. By the *double* of C, we mean the cube C^* defined by $C^* = \{c + 2x : x \in \mathbb{R}^d \text{ and } c + x \in C\}$. The double of C has the same center as C, but its edges are twice as long. The function u_d is C^∞, satisfies $0 \leq u_d \leq 1$, is identically

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

one on the standard cube $Q = \{x \in \mathbb{R}^d : 0 \le x_i \le 1, 1 \le i \le d\}$, and has support in the interior of the double $Q^\star$ of Q.

We consider an nth-order dyadic cube

$$C = 2^{-n}(a + Q), \quad a \in \mathbb{Z}^d.$$

We associate to such a cube a function u_C defined by

$$u_C(2^{-n}(a + x)) = u_d(x), \quad x \in \mathbb{R}^d.$$

This function is C^∞, satisfies $0 \le u_C \le 1$, is identically one on C, and has support in the interior of the double C^* of C.

Next, we consider an nth-order dyadic cube C in $\mathbb{T}^d$. This is the image under the projection $\pi : \mathbb{R}^d \longrightarrow \mathbb{T}^d$ of an nth-order dyadic cube $\tilde{C}$ in $\mathbb{R}^d$, i.e. $C = \pi(\tilde{C})$. We define the *double* of C to be $C^* = \pi(\tilde{C}^*)$. We define $u_C(y) = \sum_{x \in \pi^{-1}(y)} u_{\tilde{C}}(x)$. In fact, when $n \ge 1$, there is at most one $x \in \text{ supp } u_{\tilde{C}} \cap \pi^{-1}(y)$, so all but possibly one term in the sum vanishes. Obviously, u_C has the same properties as $u_{\tilde{C}}$.

If $\mathcal{F}$ is a family of nth-order dyadic cubes in $\mathbb{R}^d$ or $\mathbb{T}^d$, we define a function $u_{\mathcal{F}}$ on $\mathbb{R}^d$ or $\mathbb{T}^d$ by

$$u_{\mathcal{F}} = \frac{\sum_{\mathcal{F}} u_C}{\sum u_C}.$$

Here, the sum in the numerator is over all $C \in \mathcal{F}$. The sum in the denominator is over all nth-order dyadic cubes C in $\mathbb{R}^d$ or $\mathbb{T}^d$, according to the case we are considering. It is easy to see that if r is a non-negative integer and $0 \le \alpha \le 1$, then the $C^{r,\alpha}$ norm of $u_{\mathcal{F}}$ satisfies

$$\|u_{\mathcal{F}}\|_{r,\alpha} \le C_{r,\alpha} 2^{n(r+\alpha)}$$

where $C_{r,\alpha}$ is a positive constant that depends only on r, α, and d (and is independent of n and $\mathcal{F}$).

Next, we will construct a sequence of functions $u_0, u_1, u_2, \ldots, u_m, \ldots$ on $\mathbb{R}^d$ or on $\mathbb{T}^d$. In the case of $\mathbb{R}^d$, each of these functions will vanish in a neighborhood of Z and the common zero set of all these functions will be Z, where Z is as in §3. In the case of $\mathbb{T}^d$, the same will be true with Z replaced by $\pi(Z)$, where $\pi : \mathbb{R}^d \longrightarrow \mathbb{T}^d$ denotes the projection. By abuse of terminology, we will use Z to denote $\pi(Z)$; the two cases are distinguished by whether $Z \subset \mathbb{R}^d$ or $Z \subset \mathbb{T}^d$. (We have noted above that $\pi : Z \longrightarrow \pi(Z)$ is an isometry.) With this abuse of terminology it is possible to word the definition of the functions $u_0, u_1, u_2, \ldots$ to be the same whether the domain is $\mathbb{R}^d$ or $\mathbb{T}^d$. The definition of these functions below is meant to apply to either case.

We let $\mathcal{F}_m$ denote the family of $(n_m + \sigma)$th-order dyadic cubes (in $\mathbb{R}^d$ or $\mathbb{T}^d$, according to which case is considered) whose doubles do not meet Z. (Here, $\sigma = \sigma(d)$ is the number introduced at the end of §3.) We let $u_m = u_{\mathcal{F}_m}$. It is obvious that u_m vanishes in a neighborhood of Z and that the common zero set of all the u_m's is Z.

We will set

$$U = \sum_{m=0}^{\infty} A_m u_m,$$

for suitable positive numbers A_m. It remains to determine these numbers in such a way that the conclusions of Theorem 1 hold for this U. There are two constraints: on the one

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

hand, U should be of differentiability class $C^{2d-3,1-\epsilon}$ for every $\epsilon > 0$. On the other hand, $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ should be isometric to an interval.

In view of the estimate given above for the $C^{r,\alpha}$ norm of $u_{\mathcal{F}}$, a sufficient condition for U to be of class $C^{r,\alpha}$ is

$$\sum_{m=0}^{\infty} A_m 2^{n_m(r+\alpha)} < \infty. \qquad (\star)$$

To describe the quotient Aubry set, we need the following numbers. Let

$$b_m = \int_{E_m} \sqrt{U}\, ds, \quad c_m = \int_{F_m} \sqrt{U}\, ds,$$

where E_m is the bottom of an ultimate tooth of $C^{(m)}$, $F_m = E_m \setminus C^{(m+d)}$, and ds is the element of arc-length. Let

$$b_m^{(\infty)} = \inf(b_m, \nu_m^* b_{m+1}, \nu_m^* \nu_{m+1}^* b_{m+2}, \dots)$$

and

$$c_m^{(\infty)} = \inf(c_m, \nu_m^* c_{m+1}, \nu_m^* \nu_{m+1}^* c_{m+2}, \dots).$$

Note that $b_m^{(\infty)} = \min(b_m, \nu_m^* b_{m+1}^{(\infty)})$. Clearly, $b_m > 0$. Hence $b_m^{(\infty)} = 0$ if and only if $b_{m+1}^{(\infty)} = 0$.

PROPOSITION. *If $b_m^{(\infty)} = 0$ for some m (and therefore for every m), then $\bar{A}_{L_0,0}$ is reduced to one point. Otherwise, $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is homeomorphic to an interval. In this case, $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval if $c_m \geq \nu_m^* b_{m+1}$ for all m. It contains a sub-set isometric to an interval if there exists m_0 such that $c_m \geq \nu_m^* b_{m+1}$ for all $m \geq m_0$. In this case, if $b_{m_0-1} < \nu_{m_0-1}^* b_{m_0}^{(\infty)}$, then $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is not isometric to an interval. Moreover, if $b_m < \nu_m^* c_{m+1}^{(\infty)}$ for infinitely many m, then $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ has no sub-set isometric to an interval.*

We will finish the proof of this Proposition in §14.

In the remainder of this section, we will show as a consequence of this Proposition that it is possible to choose the numbers A_m so that U is of class $C^{2d-3,1-\epsilon}$ and any one of the four possibilities mentioned in this Proposition hold: either $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval; or it is not isometric to an interval, but contains a sub-set isometric to an interval; or it is homeomorphic to an interval, but contains no sub-set isometric to an interval; or it is reduced to a point.

Note that Theorem 1 in §2 follows from the fact that we may choose the numbers A_m so that U is of class $C^{2d-3,1-\epsilon}$ and the first possibility holds.

The first of these possibilities is the conclusion of Theorem 1 in §2. Thus, in the remainder of this section, we show that the Proposition above implies Theorem 1 in §2.

We begin by estimating various quantities. First, there exists a constant $C > 0$, independent of m, such that

$$\int_{F_m} \sqrt{u_m}\, ds = \int_{E_m} \sqrt{u_m}\, ds = C 2^{-n_m},$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

where E_m, F_m, and ds are the same as in the definition of b_m. Note that E_m is an n_mth-order dyadic edge and its length is 2^{-n_m}. The first equation above is a consequence of the fact that F_m is the set of points in E_m of distance greater than or equal to ℓ_{m+d} from the endpoints of E_m, the fact that $\ell_{m+d} \le 2^{-n_m-\sigma-3}$, and the definition of u_m. Let $\phi : E_m \longrightarrow E_{m'}$ be an affine bijection. It follows by the same reasons that $u_{m'} \circ \phi = u_m$. The second equation above is an immediate consequence.

Second, $u_{m'}|E_m = 0$, for $m' < m$, so $U|E_m = \sum_{m' \ge m} A_{m'} u_{m'}$.

Third, if A_m decreases rapidly to zero, $b_m = \int_{E_m} \sqrt{U}\, ds$ will be only a little larger than $\int_{E_m} \sqrt{A_m u_m}\, ds = C\sqrt{A_m}\, 2^{-n_m}$ when m is large; indeed

$$C\sqrt{A_m}\, 2^{-n_m} < c_m < b_m = C\sqrt{A_m}\, 2^{-n_m}\left(1 + O\left(\frac{A_{m+1} + A_{m+2} + \cdots}{A_m}\right)\right).$$

We will choose the A_m's according to the iteration formulas $A_0 = 1$ and

$$\frac{\nu_m^* \sqrt{A_{m+1}}\, 2^{-n_{m+1}}}{\sqrt{A_m}\, 2^{-n_m}} = B_m, \quad m \ge 0,$$

where the positive numbers B_m are to be determined. They will be chosen so that B_m is less than 1. Since $\nu_m^* = 2^{6+m} - 2$, the second formula may be rewritten as

$$A_{m+1}/A_m = B_m^2 2^{2(n_{m+1}-n_m)}/(2^{6+m} - 2)^2.$$

In view of the definition of n_m, we have

$$n_{m+1} - n_m - 6 - m < -m(d-1)/d + \sigma - 5 + \frac{21}{2d},$$

so A_m is rapidly decreasing for large m (provided that $B_m \le 1$ for all m), and

$$C\sqrt{A_m}\, 2^{-n_m} < c_m < b_m = C\sqrt{A_m}\, 2^{-n_m}(1 + O(2^{-m(d-1)/d})). \qquad (\dagger)$$

Since $b_m \ge C\sqrt{A_m}\, 2^{-n_m}$, in order to verify that $b_m^{(\infty)} > 0$, it is enough to verify that

$$\inf(\sqrt{A_m}\, 2^{-n_m}, \nu_m^* \sqrt{A_{m+1}}\, 2^{-n_{m+1}}, \nu_m^* \nu_{m+1}^* \sqrt{A_{m+2}}\, 2^{-n_{m+2}}, \dots) > 0.$$

In view of the iterative definition of A_m above, this will be the case if the sequence $B_0 B_1 \cdots B_m$, $m = 0, 1, \dots$, is bounded below by a positive number.

In view of ($\dagger$) and the iterative definition of A_m, we have

$$\left| \frac{\nu_m^* b_{m+1}}{c_m} - B_m \right| = O(2^{-m(d-1)/d}).$$

It is clear that we may arrange that $c_m \ge \nu_m^* b_{m+1}$ for all m if we choose the B_m's small enough. Moreover, in view of the estimate above, we may at the same time choose the B_m's so that $0 < 1 - B_m = O(2^{-m(d-1)/d})$ for large m. If we choose the B_m's in this way, the infinite product $B_1 B_2 \cdots B_m \cdots$ will converge to a positive number. Thus, according to the Proposition, we will have that $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval.

For such a choice of A_m, we have

$$\log_2 A_m = -m^2(d-1)/d \pm O(m),$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

by the iterative definition of A_m and the fact $|B_m - 1| = O(2^{-m(d-1)/d})$. Since $n_m = m^2/2d + O(m)$, it follows that $\sum A_m 2^{n_m(2d-2-\epsilon)} < \infty$. This is the condition $(\star)$ for U to be of class $C^{2d-3,1-\epsilon}$. This shows that the Proposition above implies Theorem 1 in §2.

It is easy to modify the construction above so that any one of the other three possibilities mentioned in the Proposition hold and U is of class $C^{2d-3,1-\epsilon}$.

6. *The case $L = K + U$ and $c = 0$*

In this section, we will obtain an expression for the pseudo-metric ρ_c on the Aubry set A_c in the special case when $c = 0$ and $L = K + U$, where K is the kinetic energy associated to a C^2 Riemannian metric g on a manifold M and U is a C^1 function on M. (The Lagrangian L_0 introduced in §2 is a special case of such a Lagrangian.)

Thus, for $m \in M$ and $v \in TM_m$, we have $K(v) = g(v, v)/2$ and (by abuse of notation) $U(v) = U(m)$. We will suppose that $\min U = 0$. There is no loss of generality in this, since otherwise we may replace U with $U - \min U$.

Since the Lagrangian that we are dealing with here is autonomous, we have that $\rho_c((m_0, \tau_0), (m_1, \tau_1))$ is independent of τ_0 and τ_1; we may write $\rho_c((m_0, \tau_0), (m_1, \tau_1)) = \rho_c(m_0, m_1)$ for $m_0, m_1 \in M$ and $\tau_0, \tau_1 \in \mathbb{T}$. For the case $L = K + U$ and $\min U = 0$ (as above), we have the following result.

PROPOSITION. *The pseudo-metric ρ_0 on A_0 is given by*

$$\rho_0(m_0, m_1) = \inf_\gamma 2^{3/2} \int_\gamma \sqrt{U}\, ds$$

(for $m_0, m_1 \in A_0$), where the infimum is taken over all rectifiable curves γ connecting m_0 and m_1 and ds denotes the element of arc-length associated to γ.

(Recall that a curve is said to be *rectifiable* if it is continuous and has finite arc-length. It is said to *connect* m_0 and m_1 if it is a continuous mapping $\gamma : [t_0, t_1] \longrightarrow M$ such that $\gamma(t_0) = m_0$ and $\gamma(t_1) = m_1$. Since γ is rectifiable, we may choose the parameter t so that γ is absolutely continuous, e.g. let $t = s$, the arc-length parameter. For a general absolutely continuous parameter t, the element of arc-length is defined (almost everywhere) by $ds = \|d\gamma/dt\|_2\, dt$.)

Proof. By definition,

$$\rho_0(m_0, m_1) = 2 \lim_{T \longrightarrow \infty} \left(\inf_\gamma \int_{t_0}^{t_1} L(\gamma(t), \dot\gamma(t))\, dt \right),$$

where the infimum is taken over all absolutely continuous curves $\gamma : [t_0, t_1] \longrightarrow M$ such that $\gamma(t_0) = m_0$, $\gamma(t_1) = m_1$, and $t_1 - t_0 \geq T$.

We have

$$L(\gamma(t), \dot\gamma(t))\, dt = \left(\frac{1}{2} g\left(\frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right) + U \right) dt = \left(\frac{1}{2\sigma} + U\sigma \right) ds$$

where $\sigma = dt/ds$. To minimize the integral $\int_{t_0}^{t_1} L(\gamma(t), \dot\gamma(t))\, dt$ over all parameterizations t of γ, we minimize $(1/2\sigma) + U\sigma$ pointwise. The minimum of this (as a function of σ) is

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

taken at $\sigma = 1/\sqrt{2U}$, at a point where U is positive. The minimum value is $\sqrt{2U}$. At a point where U vanishes the infimum is zero, but it is not achieved. It follows easily that the infimum of $\int_{t_0}^{t_1} L(\gamma(t), \dot\gamma(t))\,dt$, over all reparameterizations t of γ, is $\int_\gamma \sqrt{2U}\,ds$.

If $\int_\gamma dt$ is finite, i.e. γ is defined in a finite interval $[t_0, t_1]$, we may extend it to all of $\mathbb{R}$ by setting $\gamma(t) = m_0$ for $t \le t_0$ and $\gamma(t) = t_1$ for $t \ge t_1$. Since A_0 is the zero set of U and $m_0, m_1 \in A_0$, this does not change $\int L(\gamma(t), \dot\gamma(t))\,dt$. □

7. *A description of $\rho_{L_0,0}$ when $c_m \ge v_m^* b_{m+1}$*

The purpose of this section and §§8–14 is to finish the proof of the Proposition in §5.

We let L_0 be the Lagrangian defined by the formula (⋆) in §2, with U one of the functions constructed in §5. The zero set of U is $Z = A_{L_0,0} = C^{(\infty)}$. In this section, we will define a function ρ^* on $Z \times Z$. In this section and §§8–14, we will show that ρ^* is the same as the pseudo-metric $\rho_{L_0,0}$ that appears in Theorem 1 in the case that $c_m \ge v_m^* b_{m+1}$ for all m. Thus, the definition of ρ^* provides a description of $\rho_{L_0,0}$ in this case. In general, we will show that $\rho_{L_0,0} \le \rho^*$ and $\rho_{L_0,0}(x, y) > 0$ for any $x, y \in Z$ such that $\rho^*(x, y) > 0$.

It follows easily from the definitions that in the case that $c_m \ge v_m^* b_{m+1}$ for all m, ρ^* has the properties stated for $\rho_{L_0,0}$ in the Proposition in §5. This point will be discussed later in this section. In this section, we will also show that $\rho_{L_0,0} \le \rho^\star$. The proof that $\rho_{L_0,0} = \rho^\star$ in the relevant case will be postponed to later sections.

In §3, we defined

$$Z = C^{(\infty)} = \bigcup_{m \ge -1} C^{(m)} \cap C^{(m+1)},$$

where

$$C^{(-1)}(= J), C^{(0)}(= C), C^{(1)}, C^{(2)}, \ldots, C^{(m)}, \ldots$$

was a suitable sequence of iterated combs. Using the same sequence of iterated combs, we set

$$\widehat{Z} = \bigcup_{m \ge -1} C^{(m)}.$$

Thus, $Z \subset \widehat{Z}$. It follows from our definition of Z that if L is a connected component of $C^{(m)} \smallsetminus C^{(m+1)}$, $m \ge -1$, then $L \cap C^{(p)} = \phi$, for $p \ne m$, $p < m+d$, and $K = L \smallsetminus C^{(m+d)}$ is an open sub-segment of L. The association $L \mapsto K$ is a bijective correspondence between $\bigcup_{m \ge -1}$ {connected components of $C^{(m)} \smallsetminus C^{(m+1)}$} and {connected components of $\widehat{Z} \smallsetminus Z$}. Note that a connected component of $C^{(m)} \smallsetminus C^{(m+1)}$ is the bottom of an ultimate tooth of $C^{(m+1)}$.

Since Z is a Jordan arc, we may define $Z_{x,y}$ for $x, y \in Z$ to be the sub-arc of Z joining x to y. We define $\widehat{Z}_{x,y}$ to be $Z_{x,y}$ together with all connected components of $C^{(m)} \smallsetminus C^{(m+1)}$, $m \ge -1$, both of whose endpoints lie in $Z_{x,y}$. We define $Z^*_{x,y}$ to be the shortest Jordan arc in $\widehat{Z}_{x,y}$ joining x to y.

If R is an ultimate tooth of $C^{(m)}$ and ξ, η are the bottom vertices of the sides of R, we let $C_R^{(\infty)} = Z_{\xi,\eta}$.

If $x, y \in J \cap C$, then $Z^*_{x,y}$ is the line segment (in J) joining x to y. If $x \notin J \cap C$, then $x \in C_R^{(\infty)}$ for some tooth R of C. If, in addition, $y \in J \cap C$, then $Z^*_{x,y} = Z^*_{x,\xi} \cup [\xi, y]$,

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

where ξ is the endpoint of $C_R^{(\infty)}$ nearer to y and $[\xi, y]$ is the line segment (in J) joining ξ to y. If $x, y \notin J \cap C$, then $x \in C_R^{(\infty)}$ and $y \in C_{R'}^{(\infty)}$ for some teeth R, R' of C. If $R \neq R'$, then $Z_{x,y}^* = Z_{x,\xi}^* \cup [\xi, \eta] \cup Z_{\eta,y}^*$, where ξ is an endpoint of $C_R^{(\infty)}$ and η is an endpoint of $C_{R'}^{\infty}$. Moreover, these endpoints satisfy the condition that the line segment $[\xi, \eta]$ intersects R only in ξ and intersects R' only in η.

On the other hand, if $R = R'$, we choose m as large as possible so that $x, y \in C_{R^{(m)}}^{(\infty)}$, where $R^{(m)}$ is an ultimate tooth of $C^{(m)}$. Clearly, $m \geq 0$.

In this case, if (in addition) $x, y \in C^{(m)} \cap C^{(m+1)}$, then $Z_{x,y}^*$ is the sub-arc of $\partial R^{(m)}$ joining x to y. If $x \notin C^{(m)} \cap C^{(m+1)}$, then $x \in C_{R^{(m+1)}}^{(\infty)}$, where $R^{(m+1)}$ is an ultimate tooth of $C^{(m+1)}$ whose bottom lies in $\partial R^{(m)}$. If, in addition, $y \in C^{(m)} \cap C^{(m+1)}$, then $Z_{x,y}^* = Z_{x,\xi}^* \cup [\xi, y]$, where ξ is the endpoint of $C_{R^{(m+1)}}^{(\infty)}$ nearer to y in $\partial R^{(m)}$ and $[\xi, y]$ is the sub-arc of $\partial R^{(m)}$ joining ξ to y. If $x, y \notin C^{(m)} \cap C^{(m+1)}$, then $x \in C_{R^{(m+1)}}^{(\infty)}$ and $y \in C_{R^{(m+1)'}}^{(\infty)}$ for some ultimate teeth $R^{(m+1)}$ and $R^{(m+1)'}$ of $C^{(m+1)}$ whose bottoms lie in $R^{(m)}$. We have $R^{(m+1)} \neq R^{(m+1)'}$, because m was chosen as large as possible. Then $Z_{x,y}^* = Z_{x,\xi}^* \cup [\xi, \eta] \cup Z_{\eta,y}^*$, where ξ is an endpoint of $C_{R^{(m+1)}}^{(\infty)}$ and η is an endpoint of $C_{R^{(m+1)'}}^{(\infty)}$. Moreover, these endpoints satisfy the condition that the sub-arc $[\xi, \eta]$ of $\partial R^{(m)}$ intersects $R^{(m+1)}$ only in ξ and intersects $R^{(m+1)'}$ only in η.

The above descriptions of $Z_{x,y}^*$ in various cases, together with induction, provide a description of $Z_{x,y}^*$ in every case. For example, if $x \notin J \cap C$ and $y \in J \cap C$, then $Z_{x,y}^* = Z_{x,\xi}^* \cup [\xi, y]$, where $x \in C_R^{(\infty)}$ and ξ is the suitable endpoint of $C_R^{(\infty)}$ (discussed above). Then $\xi \in C_R^{(\infty)}$ so the description for the case when both points are in the same $C_{R^{(m)}}^{(\infty)}$ (in this case, with $m = 0$), applies to $Z_{x,\xi}^*$. We thus get a decomposition of $Z_{x,\xi}^*$ in at most three sub-arcs, and one of the descriptions above applies to each of these sub-arcs, this time with $m = 1$ (or the sub-arc is a sub-arc of ∂R).

Continuing in this way, we get a finite decomposition of $Z_{x,y}^*$ into possibly a sub-arc of J and into various sub-arcs of various $\partial R^{(m)}$'s. Each of these $\partial R^{(m)}$'s satisfies $m \leq \max(n, p)$, where n is the smallest integer such that $x \in C^{(n)} \cap C^{(n+1)}$ and p is the smallest integer such that $y \in C^{(p)} \cap C^{(p+1)}$.

If G is a connected component of $C^{(m)} \setminus C^{(m+1)}$ (for some $m \geq -1$) that lies in $Z_{x,y}^*$, we will call G a *gap* of $Z_{x,y}^*$. We call m the *depth* of G and denote it by $m(G)$. If there is another gap G' such that $G \subset G'$, we will call G a *secondary gap* of $Z_{x,y}^*$. Otherwise, we will call it a *primary gap* of $Z_{x,y}^*$. It is easily seen that $Z_{x,y}^*$ contains at most finitely many primary gaps. Moreover, $Z_{x,y}^*$ is contained in the union of $Z_{x,y}$ and all of the primary gaps of $Z_{x,y}^*$.

Definition. For $x, y \in Z$, we set

$$\rho^*(x, y) = 2^{3/2} \sum_G b_{m(G)}^{(\infty)},$$

where the sum is taken over all primary gaps G of $Z_{x,y}^*$.

PROPOSITION. *Let $x, y \in Z$. In the case that $c_m \geq v_m^* b_{m+1}$ for all m, we have $\rho_{L_0,0}(x, y) = \rho^*(x, y)$. In general, we have $\rho_{L_0,0}(x, y) \leq \rho^*(x, y)$ and $\rho_{L_0,0}(x, y) > 0$*

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

for any $x, y \in Z$ such that $\rho^*(x, y) > 0$.

We will give the proof in the remainder of this section and in §§8–13.

In the case that $b_m \geq \nu_m^* b_{m+1}$ for all m, the function ρ^* is a pseudo-metric on Z and the quotient metric space is isometric to an interval. Indeed, it is enough to verify that if y is between x and z on the Jordan curve Z, then $\rho^*(x, z) = \rho^*(x, y) + \rho^*(y, z)$. This is easy to verify directly from the definition of ρ^*.

In contrast, if $b_m^{(\infty)} < \nu_m^* b_{m+1}^{(\infty)}$ for some m, then ρ^* is not a pseudo-metric. Let R be an ultimate tooth of $C^{(m+1)}$ and let ξ and η be the bottom ends of the sides of R. Let $x \in C_R^{(\infty)}$. We have $\rho^*(\xi, \eta) = 2^{3/2} b_m^{(\infty)}$, $\lim_{x \to \xi} \rho^*(x, \eta) = 2^{3/2} \nu_m^* b_{m+1}^{(\infty)}$, and $\lim_{x \to \xi} \rho^*(x, \xi) = 0$. Consequently,

$$\rho^*(x, \xi) + \rho^*(\xi, \eta) < \rho^*(x, \eta),$$

for x near ξ, which shows that ρ^* is not a pseudo-metric in this case.

In the rest of this section, we show that $\rho_{L_0,0} \leq \rho^*$.

To begin, we consider an ultimate tooth R of $C^{(m+1)}$ and let ξ and η be the bottom ends of the sides of R. For $p \geq m$, we let $C_R^{(p)}$ be the segment of the Jordan arc $C^{(p)}$ joining ξ and η. It is easily seen that

$$\int_{C_R^{(p)}} \sqrt{U}\, ds = \nu_m^* \cdots \nu_p^* b_{p+1},$$

which together with the Proposition in §6 implies that

$$\rho_{L_0,0}(\xi, \eta) \leq 2^{3/2} \inf_{p \geq m} \int_{C_R^{(p)}} \sqrt{U}\, ds = 2^{3/2} b_m^{(\infty)} = \rho^*(\xi, \eta).$$

Now we consider arbitrary $x, y \in Z$. For each gap G in $Z_{x,y}$, we let $\xi(G)$ and $\eta(G)$ denote the endpoints of G. We have

$$\rho_{L_0,0}(x, y) \leq \sum_G \rho_{L_0,0}(\xi(G), \eta(G)) \leq \sum_G \rho^*(\xi, \eta) = \rho^*(x, y),$$

where the sums are taken over all primary gaps G in $Z_{x,y}^*$. The first inequality here holds because $\rho_{L_0,0}$ is pseudo-metric, the second is what we proved in the previous paragraph, and the equality follows from the definition of ρ^*.

8. *Proof of the Proposition in §7: first step*

In §3, we defined $Z \subset \mathbb{R}^d$. In §5, we also used (by abuse of notation) Z to denote $\pi(Z) \subset \mathbb{T}^d$, where $\pi : \mathbb{R}^d \longrightarrow \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is the projection, and we also used $J, C, C' \ldots$ both to denote sub-sets of $\mathbb{R}^d$ and to denote their projections on $\mathbb{T}^d$. We will continue such dual uses of notation here and introduce new sub-sets of $\mathbb{R}^d$ and denote their projections on $\mathbb{T}^d$ by the same symbols.

We let J^* be the smallest parallelepiped in $\mathbb{R}^d$ with sides parallel to the coordinate axes that contains $Z = C^{(\infty)}$. We also denote the projection of J^* on $\mathbb{T}^d$ by the same symbol. The lengths of the sides of J^* in the direction of the primary, secondary, ..., dth-ary axes of C^∞ are $\ell_0, \ell_1, \ldots, \ell_{d-1}$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Likewise, if R is an ultimate tooth of $C^{(m)}$, $m \geq 0$, we let R^* be the smallest parallelepiped in $\mathbb{R}^d$ with sides parallel to the coordinate axes that contains $C_R^{(m)}$, and denote its projection on $\mathbb{T}^d$ by the same symbol. The lengths of the sides of R^* are $\ell_{m+1}, \ldots, \ell_{m+d-1}, 2^{-n_m}$.

The *L^p-norm* on $\mathbb{R}^d$ is defined as $\|x\|_p = \left(\sum |x_i|^p\right)^{1/p-}$ if $x = (x_1, \ldots, x_d)$. The *L^∞-norm* is defined by $\|x\|_\infty = \max |x_i|$. If $\gamma : [a, b] \longrightarrow \mathbb{R}^d$ is a continuous curve, its *L^p arc-length* is defined as

$$\|\gamma\|_p = \sup \left\{ \sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|_p : a = t_0 \leq \cdots \leq t_k = b \right\}.$$

We continue to use *arc-length* without a modifier to mean L^2 arc-length.

By a rectifiable curve, we will mean one that is continuous and has finite arc-length. Obviously, the L^p arc-length of a continuous curve is finite if and only if the curve is rectifiable. If γ is rectifiable, we define the *reduced action* $RA(\gamma)$ of γ by

$$RA(\gamma) = \int_\gamma \sqrt{2U}\, ds.$$

In §6, we showed that $RA(\gamma)$ is the infimum of the actions (with respect to L_0 given by ($\star$) in §2) of all reparameterizations of γ.

LEMMA. *Suppose that $x, y \in C^{(\infty)} \subset \mathbb{T}^d$. Let γ be a rectifiable curve in $\mathbb{T}^d$ connecting x and y. Then there exists a rectifiable curve γ_1 in J^* connecting x and y such that $\|\gamma_1\|_1 \leq \|\gamma\|_1$ and $RA(\gamma_1) \leq RA(\gamma)$.*

Proof. We have $\ell_0 \leq 2^{-5}$, $\ell_1 \leq 2^{-6}$, $\ell_2 \leq 2^{-7}, \ldots$ so $\ell_0 + \cdots + \ell_{d-1} \leq 2^{-4}$ and the L^1 arc-length of any line segment in J^* is less than or equal to 2^{-4}. Moreover, U takes its maximum value $A_0 + A_1 + \cdots$ outside a suitably chosen cube Q^* of side 2^{-3}. Any curve γ that crosses the region between Q^* and $\mathbb{T}^d \setminus 2Q^*$ has L^1 arc-length greater than or equal to 2^{-3} and reduced action greater than or equal to $2^{-3}(A_0 + A_1 + \cdots)$. This will hold if γ is a continuous curve connecting x and y and does not lie entirely in $2Q^*$. For such a curve, we may reduce both the L^1 arc-length and the reduced action by replacing γ by the straight line γ_1 joining x and y.

Thus, we may assume γ lies entirely in $2Q^*$. We assume this and let $\gamma_1(t)$ be the point of J^* nearest to $\gamma(t)$. Clearly, γ_1 is continuous and $\|\gamma_1\|_1 \leq \|\gamma\|_1$. Moreover, $\|d\gamma_1(t)/dt\|_2 \leq \|d\gamma(t)/dt\|_2$ almost everywhere, so to prove that $RA(\gamma_1) \leq RA(\gamma)$, it is enough to prove that $U(\gamma_1(t)) \leq U(\gamma(t))$.

In order to prove that $U(\gamma_1(t)) \leq U(\gamma(t))$, it is enough to show that $u_m(\gamma_1(t)) \leq u_m(\gamma(t))$, for each m. For this purpose, it is more convenient to write $u_\mathcal{F}$ in the form $\Sigma_\mathcal{F} v_C$ where $v_C = u_C / \Sigma_C u_C$. Thus, in the case that $\mathcal{F}$ is a family of nth-order dyadic cubes, the denominator here is summed over all nth-order dyadic cubes.

Consider such a cube C. At a point where $\partial v_C / \partial x_i < 0$ (respectively $\partial v_C / \partial x_i > 0$), there is a cube C_1 obtained by translating C in the positive (respectively negative) direction of the x_i-axis by 2^{-n}, such that $\partial(v_C + v_{C_1})/\partial x_i = 0$. To verify that $u_m(\gamma_1(t)) \leq u_m(\gamma(t))$, we join $\gamma_1(t)$ to $\gamma(t)$ by a piecewise linear curve Λ consisting of line segments parallel to the axes.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

It follows that u_m does not decrease along Λ. Consider $C \in \mathcal{F}_m$ and a point of Λ where v_C is decreasing along Λ. We have just seen that there exists C_1 such that v_{C_1} is increasing by the same amount at the point in question. In view of the definition of $\gamma_1(t)$ in terms of $\gamma(t)$, it is easy to see that $C_1 \in \mathcal{F}_m$. In this way any decrease of v_C, with $C \in \mathcal{F}_m$, along Λ is compensated for by an equal increase of v_{C_1}, also with $C_1 \in \mathcal{F}_m$. Consequently, v_m does not decrease along Λ, so $u_m(\gamma_1(t)) \leq u(\gamma(t))$, as asserted. □

9. *Proof of the Proposition in §7: second step*

Throughout this section, we suppose that $m \geq 0$ and R is an ultimate tooth of $C^{(m)}$. If $m \geq 1$, we let r be the ultimate tooth of $C^{(m-1)}$ such that $C_R^{(\infty)} \subset C_r^{(\infty)}$. If $m = 0$, we let r be the smallest rectangle that contains C.

We let R^* be the parallelepiped associated to R in §8 (i.e. the smallest parallelepiped that contains $C_R^{(m)}$). If $m > 0$, we let r^* be the parallelepiped associated in the same way to r. If $m = 0$, we let r^* be J^*.

When $m \geq 1$, the lengths of the edges of R^* are the same as the lengths of the edges of r^* in the same direction, with two exceptions. The two exceptions are the directions of the axes of the coordinate plane parallel to the plane E that contains r. Let F be the coordinate sub-space of $\mathbb{R}^d$ spanned by the remaining $d-2$ axes. We have the product decomposition $\mathbb{R}^d = E \times F$. There is a unique parallelepiped P in F such that $r^* = r \times P$. Moreover, $R^* = (R^* \cap r) \times P$.

When $m = 0$, there is only one exceptional direction. This is the direction of the axis parallel to the line E_1 that contains J. Let F_1 denote the coordinate sub-space of $\mathbb{R}^d$ spanned by the remaining $d-1$ axes. We have the product decomposition $\mathbb{R}^d = E_1 \times F_1$. There is a unique parallelepiped P_1 in F_1 such that $r^*(= J^*) = J \times P_1$. Moreover, $R^* = (R \cap J) \times P_1$.

LEMMA. *Let $x, y \in R^*$. Let γ be a rectifiable curve in $\mathbb{T}^d$ connecting x to y such that $RA(\gamma) < \sqrt{2}\, b_m$. Then there exists a rectifiable curve γ_1 in R^* connecting x to y such that $RA(\gamma_1) \leq RA(\gamma)$ and $\|\gamma_1\|_1 \leq \|\gamma\|_1$.*

Proof. If L is a line and I is a closed interval in L, we define the *folding mapping* $F_{L,I} : L \to I$ as follows.

We partition L into closed intervals $I_i, i \in \mathbb{Z}$, all of the same length, such that $I_0 = I$, $I_i \cap I_j = \emptyset$ if $|i-j| \geq 2$, and $I_i \cap I_{i+1}$ is one point. We let $F = F_{L,I}$ be the unique continuous mapping of L onto I such that $F|I$ is the identity mapping, and $F|I_i$ is an isometry for each $i \in I$. (Thus, $F|I_i$ is orientation preserving if i is even and orientation reversing if i is odd.)

More generally, if I' is a second closed interval in L, we set $F_{I',I} = F_{L,I}|I'$. We will call this the *folding mapping* of I' into I.

Now we give the proof of the Lemma in the case that $m = 0$. We may assume that γ is in $r^* = J^*$ by the Lemma in the previous section.

We let I denote the set of points in J whose distance to $J \cap R$ is less than or equal to 2^{-n_0-1}. We let $F = F_{J,I} : J \to I$ denote the folding mapping. We let

$$\Phi = F \times \mathrm{id}_{P_1} : J^* = J \times P_1 \to I \times P_1.$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Clearly, $U \circ \Phi \leq U$.

Since γ is in J^*, we have $RA(\Phi \circ \gamma) \leq RA(\gamma)$ and $\|\Phi \circ \gamma\|_1 = \|\gamma\|_1$.

For each τ in the domain of γ, we let $\gamma_1(\tau)$ be the point in R^* nearest to $\Phi \circ \gamma(\tau)$. The same argument as at the end of the previous section shows that $U(\gamma_1(\tau)) \leq U(\Phi \circ \gamma(\tau)) \leq U(\gamma(\tau))$ and $\|d\gamma_1/d\tau\|_2 \leq \|d\Phi \circ \gamma(\tau)\ /d\tau\|_2 = \|d\gamma/d\tau\|_2$ almost everywhere, hence $RA(\gamma_1) \leq RA(\Phi \circ \gamma) \leq RA(\gamma)$. It is easy to see that $\|\gamma_1\|_1 \leq \|\Phi \circ \gamma\|_1 = \|\gamma\|_1$. This proves the Lemma in the case that $m = 0$.

In the case that $m > 0$, we use induction on m. We may assume that γ is in r^* by induction.

We divide r into two halves h and h' by a vertical bisector, i.e. a line segment joining the midpoint of the bottom of r to the midpoint of the top of r. We label the halves so that $R^* \cap r \subset h$. We let $F : r \to h$ be the mapping obtained by folding r along the vertical bisector. This is the unique continuous mapping of r into h such that $F|h$ is the identity and $F|h'$ is an isometry of h' onto h. We set $h^* = h \times P$ and $\Phi = F \times \mathrm{id}_P : r^* = r \times P \to h^* = h \times P$. We have $U \circ \Phi = U$. Hence $RA(\Phi \circ \gamma) = RA(\gamma)$ if γ is a rectifiable curve in r^*. Moreover, $\Phi \circ \gamma$ is in h^* and $\|\Phi \circ \gamma\|_1 = \|\gamma\|_1$. It follows that we may assume, without loss of generality, that γ is in h^*.

We let t denote the top of r. We let j denote the side of r that is contained in h. We let t_0 denote the image of R^* under the orthogonal projection of r^* on t. Thus, t_0 is a sub-interval of t, one of whose endpoints is $t \cap j$ and whose length is the ultimate length of $C^{(m+d-1)}$. We have seen that $r^* = j \times t \times P$. We set $r_0 = j \times t_0$ and $r_0^* = j \times t_0 \times P$. Thus, $r_0 \subset h$ and $r_0^* \subset h^*$.

If γ is a curve in h^* and $\gamma_1(t)$ is the point in r_0^* nearest to $\gamma(t)$, then $RA(\gamma_1) \leq RA(\gamma)$ and $\|\gamma_1\|_1 \leq \|\gamma\|_1$. It follows that we may assume, without loss of generality, that γ is in r_0^*.

Note that j is the side of r that meets R. We let I denote the set of points in j whose distance to $j \cap R$ is less than or equal to 2^{-n_m-1}. We let $G = G_{j,I} : j \to I$ denote the folding mapping. We let

$$\Psi = G \times \mathrm{id}_{t \times P} : r^* = j \times t \times P \to I \times t \times P.$$

If $x, y \in \mathbb{T}^d$, we set $\mathrm{dist}_p(x, y) = \min\{\|\tilde{x} - \tilde{y}\|_p : \tilde{x}, \tilde{y} \in \mathbb{R}^d \text{ and } \pi(\tilde{x}) = x, \pi(\tilde{y}) = y\}$, where $\pi : \mathbb{R}^d \to \mathbb{T}^d$ denotes the projection. If $x \in \mathbb{T}^d$ and A is a sub-set of $\mathbb{T}^d$, we set $\mathrm{dist}_p(x, A) = \inf\{\mathrm{dist}_p(x, y) : y \in A\}$. We define the *deficit zone* $DZ = DZ(r^*)$ of r^* to be $\{x \in r^* : \mathrm{dist}_\infty(x, t) \leq 3 \times 2^{-n_{m+1}-\sigma-1}\}$. For $x \in r_0^* \setminus DZ$, we have $U \circ \Psi(x) \leq U(x)$.

We have seen that we may suppose that γ is in r_0^*. We let $\gamma^*(\tau)$ be the point in R^* nearest to $\Psi \circ \gamma(\tau)$. Then $U(\gamma^*(t)) \leq U(\Psi \circ \gamma(t)) \leq U(\gamma(\tau))$, provided that $\gamma(\tau)$ in $r_0^* \setminus DZ$. Moreover, $\|d\gamma^*(\tau)/d\tau\|_p \leq \|d\gamma(\tau)/d\tau\|_p$ almost everywhere, for $1 \leq p \leq \infty$, so $RA(\gamma^*) \leq RA(\gamma)$ and $\|\gamma^*\|_1 \leq \|\gamma\|_1$. Hence, in the case that γ is in $r_0^* \setminus DZ$, we have the conclusion of the Lemma with $\gamma_1 = \gamma^*$.

It remains to consider the case when γ is in r_0^* but meets DZ. We let Σ denote the sub-set of r_0 consisting of points x such that $\mathrm{dist}(x, t) \leq 3 \times (2^{-n_m} - 2^{-n_m-\sigma-1})$. (Here, dist may stand for dist_p for any $1 \leq p \leq \infty$; it does not matter which p we choose; the set Σ is the same rectangle in r_0 in any case.) We let $\pi : r_0^* \to r_0$ be the orthogonal projection. We let $\Sigma^* = \{x \in \pi^{-1}(\Sigma) : \mathrm{dist}_\infty(x, t \cup j) \geq 3 \times 2^{-n_m-\sigma-1}\}$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

First, we consider the sub-case when the L^2 arc-length of the intersection of γ with Σ^* is greater than or equal to 2^{-n_m}. We have $U = A_m + A_{m+1} + \cdots$ on Σ^* and $U \leq A_m + A_{m+1} + \cdots$ on the bottom E_m of an ultimate tooth of $C^{(m)}$. Since the length of E_m is 2^{-n_m}, we have

$$RA(\gamma) = \sqrt{2}\int_{\gamma} \sqrt{U}\, ds \geq \sqrt{2}\int_{\gamma\cap\Sigma^*} \sqrt{U}\, ds \geq \sqrt{2}\int_{E_m} \sqrt{U}\, ds = \sqrt{2}\, b_m.$$

This contradicts the hypothesis that $RA(\gamma) < \sqrt{2}\, b_m$.

It remains to consider the sub-case when the L^2 arc-length of the intersection of γ with Σ^* is less than 2^{-n_m}. We let $[a, b]$ be the domain of γ. We let a_1 (respectively b_1) be the least (respectively greatest) $\tau \in [a, b]$ such that $\gamma(\tau) \in DZ$. We let a_2 (respectively b_2) denote the supremum (respectively infimum) of $\tau \in [a, a_1]$ (respectively $[b_1, b]$) such that $\text{dist}(\pi \circ \gamma(\tau), t) \geq 3 \times (2^{-n_m} - 2^{n_m-\sigma-1})$. We have $\text{dist}(\pi \circ \gamma(a_2), t) = \text{dist}(\pi \circ \gamma(b_2), t) = 3 \times (2^{-n_m} - 2^{-n_m-\sigma-1})$ and $\text{dist}(\pi \circ \gamma(a_1), t) \leq 3 \times 2^{-n_{m+1}-\sigma-1}$, $\text{dist}(\pi \circ \gamma(b_1), t) \leq 3 \times 2^{-n_{m+1}-\sigma-1}$. Moreover, for $a_2 \leq \tau \leq a_1$ or $b_1 \leq \tau \leq b_2$, if $\gamma(\tau) \notin \Sigma^*$, then $\text{dist}_\infty(\gamma(\tau), j) < 3 \times 2^{-n_m-\sigma-1}$. In view of our assumption that the arc-length of the intersection of γ with Σ^* is less than 2^{-n_m}, it follows that there exist a_3, a_4, b_3, b_4 with $a_2 < a_3 < a_4 < a_1$ and $b_1 < b_4 < b_3 < b_2$ such that $\text{dist}_\infty(\gamma(\tau), j) \leq 3 \times 2^{-n_m-\sigma-1}$ for $\tau = a_3, a_4, b_3, b_4$, $\text{dist}_\infty(\gamma(a_3), \gamma(a_4)) \geq 2^{-n_m}$, and $\text{dist}_\infty(\gamma(b_3), \gamma(b_4)) \geq 2^{-n_m}$.

Since $\sigma \geq \max(5, d)$ and $\text{dist}_2(\gamma(\tau), j) < \sqrt{d}\, \text{dist}_\infty(\gamma(\tau), j) \leq \sqrt{d}\, 3 \times 2^{-n_m-\sigma-1}$ for $\tau = a_3, a_4, b_3, b_4$, it is easy to see that the reduced action of curves of finite arc-length joining $\gamma(a_3)$ to $\gamma(a_4)$ (respectively $\gamma(b_4)$ to $\gamma(b_3)$) is minimized by a curve that is the concatenation of three straight lines: the straight line joining $\gamma(a_3)$ (respectively $\gamma(b_4)$) to the nearest point x_3 (respectively y_4) on j, the line segment in j joining x_3 (respectively y_4) to the point x_4 (respectively y_3) in j nearest to $\gamma(a_4)$ (respectively $\gamma(b_3)$), and the straight line joining x_4 (respectively y_3) to $\gamma(a_4)$ (respectively $\gamma(b_3)$).

Let $\gamma^* : [a, a_4] \cup [b_4, b] \to h^*$ be defined by $\gamma^*|[a, a_3] \cup [b_3, b] = \gamma|\,[a, a_3] \cup [b_3, b]$, $\gamma^*|[a_3, a_4] =$ the straight line joining $\gamma(a_3)$ to x_3, and $\gamma^*|[b_4, b_3] =$ the straight line joining y_3 to $\gamma(b_3)$. Clearly,

$$RA(\gamma^*) \leq RA(\gamma|[a, a_4] \cup [b_4, b]) \leq RA(\gamma)$$
$$\|\gamma^*\|_1 \leq \|\gamma|[a, a_4] \cup [b_4, b]\|_1 \leq \|\gamma\|_1.$$

Let $\gamma_1^* : [a, a_4] \cup [b_4, b] \to h^*$ be defined as follows: $\gamma_1^*|[a, a_2] \cup [b_2, b] = \gamma^*|[a, a_2] \cup [b_2, b]$. For $\tau \in [a_2, a_4] \cup [b_4, b_2]$, we have that, in the product decomposition $r_0^* = j \times t_0 \times P$, the j coordinate of $\gamma_1^*(\tau)$ is the point in j whose distance from $j \cap t$ is $3 \times 2^{-n_m} - 3 \times 2^{-n_m-\sigma-1}$, while the t_0 and P coordinates of $\gamma_1^*(\tau)$ are the same as the corresponding coordinates of $\gamma^*(\tau)$. Clearly $RA(\gamma_1^*) \leq RA(\gamma^*)$ and $\|\gamma_1^*\|_1 \leq \|\gamma^*\|_1$.

Note that $\gamma_1^*(a_4) = \gamma_1^*(b_4)$ which equals the point in j of distance $3 \times 2^{-n_m} - 2^{-n_m-\sigma}$ from $j \cap t$.

Consequently, if we define $\gamma_2^*(\tau) = \gamma_1^*(\tau)$, for $\tau \in [a, a_4]$ and $\gamma_2^*(\tau) = \gamma_1^*(\tau + b_4 - a_4)$, for $\tau \in [a_4, b - (b_4 - a_4)]$, we have that γ_2^* is well defined and continuous. Clearly, $RA(\gamma_2^*) = RA(\gamma_1^*)$ and $\|\gamma_2^*\|_1 = \|\gamma_1^*\|_1$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

For $a \le \tau \le b-(b_4-a_4)$, we let $\gamma_1(\tau)$ be the point on R^* nearest to $\Psi \circ \gamma_2^*(\tau)$. Since γ_2^* is in $r_0^* \setminus DZ$, we have $RA(\gamma_1) \le RA(\gamma_2^*)$. Obviously, $\|\gamma_1\|_1 \le \|\gamma_2^*\|_1$. Combining this with the equalities and inequalities already obtained, this shows that $RA(\gamma_1) \le RA(\gamma)$ and $\|\gamma_1\|_1 \le \|\gamma\|_1$. This proves the Lemma. □

10. *Proof of the Proposition in §7: third step*

Throughout this section, we suppose that $m \ge 0$ and R is an ultimate tooth of $C^{(m)}$. We let x and y be the bottom endpoints of the sides of R, and ξ and η the top endpoints, so that $[x, \xi]$ and $[y, \eta]$ are the two sides of R. We let $T = [\xi, \eta]$ denote the top of R. We divide R into two halves H and H' by a vertical bisector. We label the two halves so that $[x, \xi] \subset H$ and $[y, \eta] \subset H'$. We let E denote the plane that contains R and F the coordinate sub-space of $\mathbb{R}^d$ spanned by the remaining $(d-2)$ axes. We have the product decomposition $\mathbb{R}^d = E \times F$. There is a unique parallelepiped P in F such that $R^* = R \times P$. We let $H^* = H \times P$ and $H'^\star = H' \times P$.

LEMMA. *Let γ be a rectifiable curve in R^* connecting x to y. Then one of the following two alternatives holds:*

(a) $RA(\gamma) \ge \sqrt{2}\, c_m$;

(b) *there exists a rectifiable curve $\alpha * T * \beta$ in R^*, where α connects x to ξ in H^* and β connects η to y in H'^*, such that*

(i) $RA(\alpha * T * \beta) \le RA(\gamma)$, *and*

(ii) $\|\alpha * T * \beta\|_1 \le (1+\eta_m)\|\gamma\|_1$, *where* $\eta_m = O(2^{-m(d-1)/d})$.

Remark 1. $\alpha * T * \beta$ denotes the concatenation of α, T, and β.

Remark 2. In case (b), we will show that it is possible to take

$$\eta_m = \frac{2^{-n_m} + 2\ell_{m+1} + 3d \times 2^{-n_{m+1}-\sigma+1}}{2^{-n_m} + 2\ell_{m+1} - 3 \times 2^{-n_m-\sigma}} - 1.$$

Since $\ell_{m+1} = 2^{m+6} \times 2^{-n_{m+1}}$, the estimate for η_m given in (ii) follows.

Proof. We let $F\colon R \to H$ be the mapping obtained by folding R along the vertical bisector. We let $\Phi = F \times \mathrm{id}_P\colon R^* = R \times P \to H^* = H \times P$. We let $[t_0, t_1]$ denote the domain of γ, with $\gamma(t_0) = x$ and $\gamma(t_1) = y$. We let $t_2 = \sup\{t \in [t_0, t_1]\colon \gamma(t) \in H^*\}$. Then $\gamma(t_2) \in H^* \cap H'^*$. We define γ_1 by

$$\gamma_1(t) = \begin{cases} \Phi \circ \gamma(t), & \text{if } t_0 \le t \le t_2, \\ \gamma(t), & \text{if } t_2 \le t \le t_1. \end{cases}$$

Clearly, $RA(\gamma_1) = RA(\gamma)$ and $\|\gamma_1\|_1 = \|\gamma\|_1$. Moreover, γ_1 crosses $H^* \cap H'^*$ just once, $\gamma_1(t) \in H^*$ for $t_0 \le t \le t_2$ and $\gamma_1(t) \in H'^*$ for $t_2 \le t \le t_1$.

Thus, we can assume without loss of generality that γ crosses $H^* \cap H'^*$ just once, i.e. there exists $t_2 \in (t_0, t_1)$ such that $\gamma(t) \in H^*$ for $t \in [t_0, t_2]$ and $\gamma(t) \in H'^*$ for $t \in [t_2, t_1]$.

Let π denote the orthogonal projection of R^* on the bottom B of R. Let $B_0 = B \setminus C^{(m+d)}$. Let $[t_3, t_4]$ be the connected component of $(\pi \circ \gamma)^{-1}(B_0)$ that contains t_2. We define the *enlarged deficit zone* $\mathrm{EDZ} = \mathrm{EDZ}(R^*)$ of R^* to be $\{x \in R^*\colon \mathrm{dist}_\infty(x, T) \le 3 \times 2^{-n_m-\sigma-1}\}$. There are two cases depending on whether $\gamma|[t_3, t_4]$ meets EDZ or not.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

First, we consider the case when $\gamma|[t_3, t_4]$ does not meet EDZ. For $t \in [t_3, t_4]$, we have $\gamma(t) \in \pi^{-1}(B_0) \setminus \text{EDZ}$ and it follows that $U(\gamma(t)) \geq U(\pi(\gamma(t))$. Hence

$$RA(\gamma) \geq \int_{t_3}^{t_4} \sqrt{2(U \circ \gamma)}\, ds \geq \int_{B_0} \sqrt{2U}\, ds = \sqrt{2}\, c_m.$$

Thus, case (a) of the conclusion holds.

Now, we consider the case when $\gamma|[t_3, t_4]$ meets EDZ. We let $\pi_1 \colon R^* \to [x, \xi]$ and $\pi_2 \colon R^* \to [y, \eta]$ denote the orthogonal projections. We let $t_5 = \sup\{t \in [t_0, t_3] \colon \text{dist}_\infty (\gamma(t), [x, \xi]) \leq 3 \times 2^{-n_{m+1}-\sigma-1}$ and $\text{dist}(\pi_1(\gamma(t)), \xi) \geq 2^{-n_{m+1}}\}$ and $t_6 = \inf\{t \in [t_4, t_1] \colon \text{dist}_\infty (\gamma(t), [y, \eta]) \leq 3 \times 2^{-n_{m+1}-\sigma-1}$ and $\text{dist}\,(\pi_2(\gamma(t)), \eta) \geq 2^{-n_{m+1}}\}$. Let $\alpha = \gamma|[t_0, t_5] * [\gamma(t_5), \pi_1\gamma(t_5)] * [\pi_1\gamma(t_5), \xi]$ and $\beta = [\eta, \pi_2\gamma(t_6)] * [\pi_2\gamma(t_6), \gamma(t_6)] * \gamma|[t_6, t_1]$.

To compare $\|\gamma\|_1$ and $\|\alpha * T * \beta\|_1$, we express these quantities in the form

$$\|\gamma\|_1 = A + B + C$$

and

$$\|\alpha * T * \beta\|_1 = A + B^* + C,$$

where

$$\begin{aligned} A &= \|\gamma|[t_0, t_5]\|_1, \\ B &= \|\gamma|[t_5, t_6]\|_1, \\ C &= \|\gamma|[t_6, t_1]\|_1, \end{aligned}$$

and

$$B^* = \|\gamma(t_5) - \pi_1\gamma(t_5)\|_1 + \|\xi - \pi_1\gamma(t_5)\|_1 + 2^{-n_m} + \|\eta - \pi_2\gamma(t_6)\|_1 + \|\gamma(t_6) - \pi_2\gamma(t_6)\|_1.$$

Clearly, $A \geq \widehat{A}$, $B \geq \widehat{B}$, and $C \geq \widehat{C}$, where

$$\begin{aligned} \widehat{A} &= \ell_{m+1} - \|\xi - \pi_1\gamma(t_5)\|_1 + \|\gamma(t_5) - \pi_1\gamma(t_5)\|_1, \\ \widehat{B} &= \|\xi - \pi_1\gamma(t_5)\|_1 + \|\eta - \pi_2\gamma(t_6)\|_1 - 3 \times 2^{-n_m-\sigma} \\ &\quad + 2^{-n_m} - \|\gamma(t_5) - \pi_1\gamma(t_5)\|_1 - \|\gamma(t_6) - \pi_2\gamma(t_6)\|_1, \\ \widehat{C} &= \ell_{m+1} - \|\eta - \pi_2\gamma(t_6)\|_1 + \|\gamma(t_6) - \pi_2\gamma(t_6)\|_1. \end{aligned}$$

Hence

$$\begin{aligned} \frac{\|\alpha * T * \beta\|_1}{\|\gamma\|_1} &= \frac{A + B^* + C}{A + B + C} \leq \frac{\widehat{A} + B^* + \widehat{C}}{\widehat{A} + B + \widehat{C}} \leq \frac{\widehat{A} + B^* + \widehat{C}}{\widehat{A} + \widehat{B} + \widehat{C}} \\ &= \frac{2\ell_{m+1} + 2^{-n_m} + 2(\|\gamma(t_5) - \pi_1\gamma(t_5)\|_1 + \|\gamma(t_6) - \pi_2\gamma(t_6)\|_1)}{2\ell_{m+1} + 2^{-n_m} - 3 \times 2^{-n_m-\sigma}}. \end{aligned}$$

Since $\|\gamma(t_5) - \pi_1\gamma(t_5)\|_1 \leq 3d \times 2^{-n_{m+1}-\sigma-1}$ and $\|\gamma(t_6) - \pi_1\gamma(t_6)\|_1 \leq 3d \times 2^{-n_{m+1}-\sigma-1}$, we have proved the estimate for η_m given in Remark 2. As pointed out there, (ii) in case (b) of the Lemma follows.

It remains to prove (i). We let ξ_0 (respectively η_0) denote the bottom endpoint of the connected component of $[x, \xi] \cap C^{(m+1)}$ (respectively $[y, \eta] \cap C^{(m+1)}$) that contains ξ

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

(respectively η). Thus, $\text{dist}(\xi_0, \xi) = 3 \times 2^{-n_{m+1}}$. We set $J_0 = [\pi_1\gamma(t_5), \xi_0]$ if $\pi_1\gamma(t_5) \notin [\xi_0, \xi]$ and $J_0 = \phi$ if $\pi_1\gamma(t_5) \in [\xi_0, \xi]$. We set $J_1 = [\eta_0, \pi_2\gamma(t_6)]$ if $\pi_2\gamma(t_6) \notin [\eta, \eta_0]$ and $J_1 = \phi$ if $\pi_2\gamma(t_6) \in [\eta, \eta_0]$. The function U vanishes identically on $[\xi_0, \xi] \cup T \cup [\eta, \eta_0]$. Hence

$$RA(\alpha * T * \beta) = RA(\gamma|[t_0, t_5]) + RA([\gamma(t_5), \pi_1\gamma(t_5)]) + RA(J_0) + RA(J_1) + RA([\pi_2\gamma(t_6), \gamma(t_6)) + RA(\gamma|[t_6, t_1]).$$

Thus, to prove (i), it is enough to prove that

$$RA(J_0) + RA(J_1) + RA([\gamma(t_5), \pi_1\gamma(t_5)]) + RA([\pi_2\gamma(t_6), \gamma(t_6)]) \leq RA(\gamma|[t_5, t_6]). \quad (\star)$$

The function U vanishes identically on $J_0 \cap C^{(m+1)}$. Each component of $J_0 \setminus C^{(m+1)}$ is an n_{m+1}th-order dyadic edge E_{m+1}, or part of such an edge (in which case the endpoint $\pi_1\gamma(t_5)$ of J_0 is in the edge). This dyadic edge is the bottom of an ultimate tooth of $C^{(m+1)}$. Let E'_{m+1} be the n_{m+1}th-order dyadic edge immediately above E_{m+1} on $[x, \xi]$. Let $t_7 \in [t_3, t_4]$ minimize $\text{dist}_\infty(\gamma(t), T)$ over all $t \in [t_3, t_4]$. If $\pi_1\gamma(t_7)$ is above E'_{m+1} on $[x, \xi]$, then there exist $t^*, \hat{t} \in [t_5, t_7]$ such that $\pi_1\gamma(t^*)$ equals the bottom endpoint of E'_{m+1}, $\pi_1\gamma(\hat{t})$ equals the top endpoint of E'_{m+1}, and $\pi_1\gamma(t) \in E'_{m+1}$ for all $t \in [t^*, \hat{t}]$. Moreover, $\text{dist}_\infty(\gamma(t), [x, \xi]) \geq 3 \times 2^{-n_{m+1}-\sigma-1}$ in view of the definition of t_5. It follows that $U(\gamma(t)) \geq U(\tau)$, where τ is the image of $\pi_1\gamma(t)$ under the orientation-preserving isometry of E'_{m+1} onto E_{m+1}. Hence,

$$\sqrt{2}\, b_{m+1} = RA(E_{m+1}) \leq RA(\gamma|[t^*, \hat{t}]).$$

In this way, we have found a portion of the curve $\gamma|[t_5, t_7]$ whose reduced action dominates the reduced action of E_{m+1}, provided that $\pi_1\gamma(t_7)$ is above E'_{m+1}. For different components of $J_0 \setminus C^{(m+1)}$, the corresponding portions of γ are mutually disjoint, since their projections on J_0 are mutually disjoint. In the same way, we may find, for an appropriate component of $J_1 \setminus C^{(m+1)}$, a portion of the curve $\gamma|[t_7, t_6]$ whose reduced action dominates the reduced action of that component. Moreover, for different components, the corresponding portions are mutually disjoint.

Now we consider two sub-cases depending on whether $\text{dist}_\infty(\gamma(t_7), T) \leq 3 \times 2^{-n_{m+1}-\sigma-1}$ or not.

First, we consider the sub-case when $\text{dist}_\infty(\gamma(t_7), T) \geq 3 \times 2^{-n_{m+1}-\sigma-1}$. We let $\widehat{B}_0 = \{x \in B_0 \colon \text{dist}(x, \partial B_0) \geq 3 \times 2^{-n_{m+1}-\sigma-1}\}$. We let $[\hat{t}_3, \hat{t}_4]$ be the connected component of $(\pi \circ \gamma)^{-1}(\widehat{B}_0)$ that contains t_2. For $t \in [\hat{t}_3, \hat{t}_4]$, we have $u_j(\gamma(t)) = 1$ if $j \geq m + 1$. This implies that $U(\gamma(t)) \geq U(\tau)$ for any $\tau \in J_0 \cup J_1 \cup [\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$. Hence, the reduced action of any portion of $J_0 \cup J_1 \cup [\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$ is dominated by the reduced action of any portion of $\gamma|[\hat{t}_3, \hat{t}_4]$ of the same length.

Thus, to complete the proof of (i) in the sub-case that we are considering, it is enough to show that the length of the part of $J_0 \cup J_1 \cup [\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$ not considered above is not greater than the length of the admissible part of $\gamma|[\hat{t}_3, \hat{t}_4]$, i.e. the part of $\gamma|[\hat{t}_3, \hat{t}_4]$ not used above.

The part of $J_0 \cup J_1 \cup [\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$ not considered above consists of $[\gamma(t_5), \pi_1\gamma(t_5)]$, $[\pi_2\gamma(t_6), \gamma(t_6)]$, and those n_{m+1}th-order dyadic edges E_{m+1} on

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

$J_0 \setminus C^{(m+1)}$ or $J_1 \setminus C^{(m+1)}$ such that $\pi_1\gamma(t_7)$ is not above the dyadic edge E'_{m+1} immediately above E_{m+1}.

The length of each of $[\gamma(t_5), \pi_1\gamma(t_5)]$ and $[\pi_2\gamma(t_6), \gamma(t_6)]$ is less than or equal to $3\sqrt{d} \times 2^{-n_{m+1}-\sigma-1}$, so the sum of their lengths is less than or equal to $2^{-n_{m+1}}$. Since $\gamma(t_7) \in \mathrm{EDZ}$, we have $\mathrm{dist}_\infty(\gamma(t_7), T) \leq 3 \times 2^{-n_m-\sigma-1}$. It follows that the sum of the lengths of all the dyadic edges not considered above is less than or equal to $3 \times 2^{-n_m-\sigma-1}$. Thus, the part not considered above has total length less than or equal to $3 \times 2^{-n_m-\sigma-1} + 2^{-n_{m+1}}$.

To obtain a lower bound on the length of the admissible part of $\gamma|[\hat{t}_3, \hat{t}_4]$, we begin by obtaining a lower bound on the length of $\gamma|[\hat{t}_3, \hat{t}_4]$. We let $v_1 = \mathrm{dist}(\pi_1\gamma(\hat{t}_3), \pi_1\gamma(t_7))$, $v_2 = \mathrm{dist}(\pi_2\gamma(t_7), \pi_2\gamma(\hat{t}_4))$, $h_1 = \mathrm{dist}(\pi\gamma(\hat{t}_3), \pi\gamma(t_7))$, and $h_2 = \mathrm{dist}(\pi\gamma(t_7), \pi\gamma(\hat{t}_4))$. The length of $\gamma|[\hat{t}_3, t_7]$ is not less than the Euclidean distance between $\gamma(\hat{t}_3)$ and $\gamma(t_7)$. By the Pythagorean Theorem, this is greater than or equal to $\sqrt{v_1^2 + h_1^2}$. Likewise, the length of $\gamma|[t_7, \hat{t}_4]$ is greater than or equal to $\sqrt{v_2^2 + h_2^2}$. Hence, the length of $\gamma[\hat{t}_3, \hat{t}_4]$ is greater than or equal to $\sqrt{v_1^2 + h_1^2} + \sqrt{v_2^2 + h_2^2}$.

The part of $\gamma|[\hat{t}_3, t_7]$ used to dominate part of J_0 is used to dominate those E_{m+1} such that $\pi_1\gamma(t_7)$ is not above E'_{m+1}. The sum of the lengths of all such E_{m+1} is less than or equal to $v_1/2 + 2^{-n_{m+1}}$. Since $U(\gamma(t)) \geq U(\tau)$, for $t \in [\hat{t}_3, t_7]$ and $\tau \in E_{m+1}$, it is enough to use a part of $\gamma|[\hat{t}_3, t_7]$ whose length is the length of E_m to dominate E_m. Thus, it is enough to use a part of $\gamma|[\hat{t}_3, t_7]$ of length less than or equal to $v_1/2 + 2^{-n_{m+1}}$. Similar considerations for J_1 in place of J_0 lead to the conclusion that the admissible part of $\gamma|[\hat{t}_3, \hat{t}_4]$ has length greater than or equal to $\sqrt{v_1^2 + h_1^2} + \sqrt{v_2^2 + h_2^2} - v_1/2 - v_2/2 - 2^{-n_{m+1}+1} \geq (\sqrt{3}/2)h - 2^{-n_{m+1}+1}$, where $h = h_1 + h_2 = 2^{-n_m} - 2\ell_{m+d+1} - 3 \times 2^{-n_{m+1}-\sigma-1}$.

Thus, the verification of (i) in the sub-case that we are considering comes down to checking

$$3 \times 2^{-n_m-\sigma-1} + 2^{-n_{m+1}} \leq (\sqrt{3}/2)(2^{-n_m} - 2\ell_{m+d+1} - 3 \times 2^{-n_{m+1}-\sigma-1}) - 2^{-n_{m+1}+1}.$$

This is easily verified.

Now, we consider the other sub-case, viz. $\mathrm{dist}_\infty(\gamma(t_7), T) \leq 3 \times 2^{-n_{m+1}-\sigma-1}$. In this sub-case, for each component E_{m+1} of $J_0 \setminus C^{(m+1)}$, we have that $\pi_1\gamma(t_7)$ is above the n_{m+1}th-order dyadic edge E'_{m+1} immediately above E_{m+1} on $[x, \xi]$. For this reason, the discussion above provides a portion of the curve $\gamma|[t_5, t_7]$ whose reduced action dominates the reduced action of E_{m+1}. The same discussion applies if E_{m+1} is a component of $J_1 \setminus C^{(m+1)}$.

All that remains is to find an unused portion of the curve $\gamma|[t_5, t_7]$ whose reduced action dominates the reduced action of $[\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$. For this, we may use $\pi_1^{-1}(\widehat{E}_{m+1})$, where $\widehat{E}_{m+1}$ is the second n_{m+1}th-order dyadic edge beneath ξ. We have $u_j(\gamma(t)) = 1$ for t in this set and the length of this segment of γ is greater than or equal to $2 \times 2^{-n_{m+1}} \geq 3\sqrt{d} \times 2^{-n_{m+1}-\sigma}$ which is greater than or equal to the length of $[\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$. This shows that the reduced action of $\pi^{-1}(\widehat{E}_{m+1})$ dominates the reduced action of $[\gamma(t_5), \pi_1\gamma(t_5)] \cup [\pi_2\gamma(t_6), \gamma(t_6)]$. □

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

11. *Proof of the Proposition in §7: fourth step*

Throughout this section, we let r be an ultimate tooth of $C^{(m)}$ for some $m \geq 0$ or we let $r = J$ and $m = -1$. We let x and ξ be the endpoints of a side of r (in the case that $m \geq 0$) or the endpoints of J (in the case that $m = -1$). We let $Y_{x,\xi}$ denote the union of $[x, \xi]$ and all R^* such that R is an ultimate tooth of $C^{(m+1)}$ whose bottom lies in $[x, \xi]$. Here, R^* is the parallelepiped associated to R in §8, i.e. the smallest parallelepiped that contains $C_R^{(\infty)}$. Note that $Z_{x,\xi} \subset Y_{x,\xi}$.

LEMMA. *Let $y, z \in Z_{x,\xi}$. Let γ be a rectifiable curve in $\mathbb{T}^d$ connecting y to z. Then there exists a rectifiable curve γ_1 in $Y_{x,\xi}$ connecting y to z such that $RA(\gamma_1) \leq RA(\gamma)$, provided that at least one of the following two conditions is satisfied:*

(a) *$c_{\widetilde{m}} \geq \nu^*_{\widetilde{m}} b_{\widetilde{m}+1}$ for all $\widetilde{m} \geq m$;*

(b) *$y, z \in [x, \xi]$ and γ is in r^*.*

Moreover, if (b) is satisfied, then it is possible to choose γ_1 so that not only $RA(\gamma_1) \leq RA(\gamma_1)$, but also $\|\gamma_1\|_1 \leq (1 + \eta^*_m)\, \|\gamma\|_1$, where $\eta^*_m = O(2^{-m(d-1)/d})$.

Proof. First, we consider the case when (a) holds.

For p large enough, $y, z \in C^{(p)}$. Let $\tilde{\gamma}$ be the arc of $C^{(p)}$ between y and z. Since $c_{\widetilde{m}} \geq \nu^*_{\widetilde{m}} b_{\widetilde{m}+1}$ for all $\widetilde{m} \geq m$, it follows that $RA(\tilde{\gamma}) \leq c_m/\sqrt{2}$. (Compare with the proof that $\rho_{l_0,0}(\xi, \eta) \leq \rho^*(\xi, \eta)$ at the end of §7.) Thus, we may assume, without loss of generality, that $RA(\gamma) \leq c_m/\sqrt{2}$. We may then apply the Lemma in §9 and conclude that there exists a rectifiable curve γ_1 in r^* connecting y to z such that $RA(\gamma_1) \leq RA(\gamma)$. Thus, we may assume that γ is in r^*(or J^* in the case $m = -1$), in the case when (a) holds. This is true by assumption in the case when (b) holds.

The next part of the argument depends on the assumption that γ is in r^*, but does not otherwise use either hypothesis (a) or hypothesis (b).

We divide r into two halves h and h' by a vertical bisector. We label the halves so that $[x, \xi] \subset h$. We let E denote the plane that contains r and F the coordinate sub-space of $\mathbb{R}^d$ that contains the $d-2$ coordinate axes that are not parallel to lines in E. Then $\mathbb{R}^d = E \times F$ and there is a unique parallelepiped P in F such that $r^* = r \times P$. As in §8, we let $h^* = h \times P$. We let $F : r \to h$ be the mapping obtained by folding r along the vertical bisector and $\Phi = F \times \mathrm{id}_P : r^* \to h^*$. Just as in §8, we have $RA(\Phi \circ \gamma) = RA(\gamma)$ and $\|\Phi \circ \gamma\|_1 = \|\gamma\|_1$, if γ is a rectifiable curve in r^*. Thus, we may assume, without loss of generality, that γ is in h^*.

We let t denote the top of r. We let t_0 denote the image of R^* under the orthogonal projection of r^* on t, where R is an ultimate tooth of $C^{(m+1)}$ whose bottom is in $[x, \xi]$ and R^* is the associated parallelepiped. Thus, t_0 is a sub-interval of t, one of whose endpoints is ξ and whose length is the ultimate length of $C^{(m+d)}$. We set $r_0 = [x, \xi] \times t_0$ and $r_0^* = r_0 \times P$. Thus $r_0 \subset h$ and $r_0^* \subset h^*$.

If γ is a curve in h^* and $\gamma_1(t)$ is the point in r_0^* nearest to $\gamma(t)$, then $RA(\gamma_1) \leq RA(\gamma)$ and $\|\gamma_1\|_1 \leq \|\gamma\|_1$. It follows that we may assume, without loss of generality, that γ is in r_0^*.

In the case that $m = -1$, this part of the argument is not needed; in this case, we substitute J^* for r_0^* in what follows.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

We let $\pi_1 : r_0^* \to [x,\xi]$ denote the orthogonal projection. For each connected component e of $[x,\xi]\cap C^{(m+1)}$ that is contained in $Z_{y,z}$, we choose $t_2 = t_2^e$ and $t_3 = t_3^e$ so that $t_0 < t_2 < t_3 < t_1$, where $[t_0,t_1]$ is the domain of γ, so that $\pi_1\gamma([t_2,t_3]) = e$, and so that $\pi_1\gamma(t_2)$ and $\pi_1\gamma(t_3)$ are the endpoints of e.

We begin by considering the case when, for each e as above, we have $RA(\gamma|[t_2^e,t_3^e]) \geq \sqrt{2}\,\tilde{b}_{m+1}$, where $\tilde{b}_{m+1} = \int_{G_{m+1}} \sqrt{U}\,ds$, and G_{m+1} is the line segment $e\times\eta \subset r_0 = [x,\xi]\times t_0$ with η the endpoint of t_0 other than ξ.

Since $Z_{x,\xi}$ is a Jordan arc, it has a natural order, with x before ξ. Without loss of generality, we may suppose that y is before z in this order. We let $\tilde{e}$ denote the first e as above after y, with respect to this order. We let $\widetilde{E}$ denote the closure of the connected component of $[x,\xi]\setminus C^{(m+1)}$ immediately below $\tilde{e}$, where we take x to be the bottom and ξ the top of $[x,\xi]$. We let $\tilde{\zeta} = \widetilde{E}\cap\tilde{e}$, i.e. $\tilde{\zeta}$ is the bottom endpoint of $\tilde{e}$ and the top endpoint of $\widetilde{E}$ along $[x,\xi]$. We let $\widetilde{R}$ be the ultimate tooth of $C^{(m+1)}$ whose bottom is $\widetilde{E}$. We let $[\tilde{\zeta},\widetilde{\omega}]$ be the side of $\widetilde{R}$ whose bottom is $\tilde{\zeta}$ (and whose top is $\widetilde{\omega}$).

We let $\hat{e}$ denote the last e as above before z with respect to the order above on $Z_{x,\xi}$. We let $\widehat{E}$ denote the closure of the connected component of $[x,\xi]\setminus C^{(m+1)}$ immediately above $\hat{e}$ on $[x,\xi]$. We let $\hat{\zeta} = \widehat{E}\cap\hat{e}$. We let $\widehat{R}$ be the ultimate tooth of $C^{(m+1)}$ whose bottom is $\widehat{E}$. We let $[\hat{\zeta},\widehat{\omega}]$ be the side of $\widehat{R}$ whose bottom is $\hat{\zeta}$ (and whose top is $\widehat{\omega}$).

We will take γ_1 to be of the form $\gamma_1 = A * [\tilde{\zeta},\hat{\zeta}] * B$, where A is a suitable rectifiable curve joining y to $\tilde{\zeta}$ and B is a suitable rectifiable curve connecting $\hat{\zeta}$ to B.

We observe that $RA([\tilde{\zeta},\hat{\zeta}])$ is $\sqrt{2}\,b_{m+1}\times$ number of connected components of $[\tilde{\zeta},\hat{\zeta}]\setminus C^{(m+1)}$. The number of connected components e of $[x,\xi]\cap C^{(m+1)}$ satisfying $e\subset Z_{y,z}$ is one greater than the number of connected components of $[\tilde{\zeta},\hat{\zeta}]\setminus C^{(m+1)}$. In the case that we are considering, $RA(\gamma|[t_2^e,t_3^e]) \geq \sqrt{2}\,\tilde{b}_{m+1} > \sqrt{2}\,b_{m+1}$ for each such e. Thus, we may use these reduced actions to dominate $RA([\tilde{\zeta},\hat{\zeta}])$ and have at least $\sqrt{2}\,\tilde{b}_{m+1}$ left over. Thus, in order to prove that $RA(\gamma_1) \leq RA(\gamma)$, it will be enough to show that

$$RA(A) \leq RA(\gamma|[t_0,t_2^{\tilde{e}}]) + \tilde{b}_{m+1}/\sqrt{2}$$

and

$$RA(B) \leq RA(\gamma|[t_3^{\hat{e}},t_1]) + \tilde{b}_{m+1}/\sqrt{2}.$$

Thus, the problem of choosing γ_1 so that $RA(\gamma_1) \leq RA(\gamma)$ reduces to the problem of choosing A and B so the inequalities above are satisfied. We will show the existence of such a curve A; the same reasoning shows the existence of B.

If $RA(\gamma|[t_0,t_2^{\tilde{e}}]) \geq \sqrt{2}\,c_{m+1}$, then we may set $A = Z^*_{y,\tilde{\zeta}}$, where the latter is defined as in §7, i.e. $Z^*_{y,\tilde{\zeta}}$ is the shortest Jordan arc in $\widehat{Z}_{y,\tilde{\zeta}}$ joining y to $\tilde{\zeta}$. In the case that $y\in[x,\xi]$, we have that $Z^*_{y,\zeta}$ is the straight line segment joining y to ζ. In this case, $RA(A) = \sqrt{2}\,b_{m+1} \leq \sqrt{2}\,c_{m+1} + \tilde{b}_{m+1}/\sqrt{2}$. In the case that $y\notin[x,\xi]$, we have that $Z^*_{y,\tilde{\zeta}}\subset R^*$. In this case, hypothesis (b) does not hold, so we may assume that hypothesis (a) holds. This, together with the argument at the end of §7 to show that $\rho_{L_0,0}\leq\rho^*$, implies that $RA(A)\leq\sqrt{2}\,c_{m+1}$. Thus we have the desired inequality in the case that $RA(\gamma|[t_0,t_2^{\tilde{e}}]) \geq \sqrt{2}\,c_{m+1}$.

Now we consider the case when $RA(\gamma|[t_0,t_2^{\tilde{e}}]) < \sqrt{2}\,c_{m+1}$. We let $\tilde{\alpha}$ be the point on the side of $\widetilde{R}$ opposite to $[\tilde{\zeta},\widetilde{\omega}]$ at a distance of $2^{-n_{m+1}}$ from the top of $\tilde{R}$. If $y\in Z_{\tilde{\alpha},\tilde{\zeta}}$,

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

we let $A = Z^*_{y,\tilde\zeta}$. In this case, hypothesis (b) does not hold, so we may use hypothesis (a), together with the argument at the end of §7. This time, we obtain that $RA(A) \le c_{m+1}/\sqrt{2} < \tilde b_{m+1}/\sqrt{2}$, so we have the desired inequality in the case that $y \in Z_{\tilde\alpha,\tilde\zeta}$.

In the case that $y \notin Z_{\tilde\alpha,\tilde\zeta}$, we use a construction similar to that in the previous section. We continue to let π_1 denote the orthogonal projection of r_0^* on $[x,\xi]$. We let $\tilde\beta$ and $\tilde\delta$ be the endpoints of $\widetilde E \smallsetminus C^{(m+d+1)}$ with $\tilde\beta$ nearer to $\tilde\zeta$. (Thus $\operatorname{dist}(\tilde\beta,\tilde\zeta) = \ell_{m+d+1}$.) We let $t_4 = \inf\{t \ge t_0 : \pi_1\gamma(t) = \tilde\beta\}$ and $t_5 = \sup\{t \le t_4 : \pi_1\gamma(t) = \tilde\delta\}$. Since $RA(\gamma|[t_5,t_4]) \le RA(\gamma|[t_0,t_2^{\tilde e}]) < \sqrt{2}\,c_{m+1}$, we have that $\gamma|[t_5,t_4]$ passes through the extended danger zone $\mathrm{EDZ} = \mathrm{EDZ}(\tilde R^*)$ of $\tilde R^*$, i.e. $\{x \in \tilde R^* : \operatorname{dist}_\infty(x,\tilde T) \le 3\times 2^{-n_{m+1}-\sigma-1}\}$, where $\tilde T$ denotes the top of $\tilde R$. This follows from the discussion in the previous section.

We let $\tilde\epsilon$ be the endpoint of $\widetilde E$ different from $\tilde\zeta$. (Thus, for the natural order on $[x,\xi]$ with $x < \xi$ we have $x < \tilde\epsilon < \tilde\delta < \tilde\beta < \tilde\zeta < \hat\zeta < \xi$.) We let π_2 be the orthogonal projection of r_0^* on $\pi_1^{-1}(\tilde\epsilon)$. For $t_0 \le t \le t_4$, let $\gamma_2(t) = \pi_2\gamma(t)$ when $\pi_1\gamma(t) \le \tilde\epsilon$, and $\gamma_2(t) = \gamma(t)$ when $\pi_1\gamma(t) \ge \tilde\epsilon$. We let $\tilde\eta$ be the top endpoint of the side of $\widetilde R$ of which $\tilde\epsilon$ is the bottom endpoint. (Thus, the two sides of $\widetilde R$ are $[\tilde\zeta,\widetilde\omega]$ and $[\tilde\epsilon,\tilde\eta]$.) We let $t_6 = \sup\{t \in [t_0,t_4] : \operatorname{dist}_\infty(\gamma_2(t),[\tilde\epsilon,\tilde\eta]) \le 3\times 2^{-n_{m+2}-\sigma-1}$ and $\operatorname{dist}(\pi_2\gamma(t),\hat\eta) \ge 3\times 2^{-n_{m+1}-\sigma-1}\}$. We let $\widetilde\gamma_3 = \gamma_2|[t_0,t_6]*[\gamma_2(t_6),\pi_2\gamma(t_6)]*[\pi_2\gamma(t_6),\tilde\epsilon]$. The estimates of the previous section show that $RA(\widetilde\gamma_3) \le RA(\gamma|[t_0,t_4]) \le RA(\gamma|[t_0,t_2^{\tilde e}])$.

If $y \notin [x,\xi]$ then $y \in \tilde R^*$ and hypothesis (a) holds. We set $A = \widetilde\gamma_3 \star \widetilde T \star [\widetilde\omega,\tilde\zeta]$. Here, $\widetilde T = [\tilde\eta,\widetilde\omega]$ is the top of $\widetilde R$. We have $RA(\widetilde T) = 0$ and $RA([\widetilde\omega,\tilde\zeta]) = \nu^*_{m+1}b_{m+2}/\sqrt{2} \le c_{m+1}/\sqrt{2} < \tilde b_{m+1}/\sqrt{2}$. Hence $RA(A) \le RA(\gamma|[t_0,t_2^{\hat e}]) + \tilde b_{m+1}/\sqrt{2}$, as desired.

In the case that $y \in [x,\xi]$, we consider two sub-cases, depending on which is larger, $RA(\gamma|[t_0,t_2^{\tilde e}])$ or $\tilde b_{m+1}/\sqrt{2}$. First, we consider the sub-case when $RA(\gamma|[t_0,t_2^{\tilde e}]) \le \tilde b_{m+1}/\sqrt{2}$. We let $\widetilde\gamma_4$ be the mirror image of $\widetilde\gamma_3^{-1}$ in $\widetilde R^*$. Here $\widetilde\gamma_3^{-1}$ means $\widetilde\gamma_3$ traced out in the opposite direction. Since $\widetilde\gamma_3$ is a rectifiable curve in $\widetilde R^*$ connecting $\tilde\epsilon$ to $\tilde\eta$, we have that $\widetilde\gamma_4$ is a rectifiable curve in $\widetilde R^*$ connecting $\widetilde\omega$ to $\tilde\zeta$ and $RA(\widetilde\gamma_4) = RA(\widetilde\gamma_3)$. We set $A = [y,\tilde\epsilon] \star \widetilde\gamma_3 \star T \star \widetilde\gamma_4$. Note that $y \in [x,\tilde\epsilon]$ and $RA([y,\tilde\epsilon]) = 0$. Hence $RA(A) = RA(\widetilde\gamma_3) + RA(\widetilde\gamma_4) = 2RA(\widetilde\gamma_3) \le 2RA(\gamma|[t_0,t_2^{\tilde e}]) \le RA(\gamma|t_0,t_2^{\tilde e}]) + \tilde b_{m+1}/\sqrt{2}$, which is the desired inequality.

Now, we consider the other sub-case, i.e. $RA(\gamma|[t_0,t_2^{\tilde e}]) \ge \tilde b_{m+1}/\sqrt{2}$. In this case, we set $A = [y,\tilde\zeta]$. We have $RA(A) = \sqrt{2}\,b_{m+1} < \sqrt{2}\,\tilde b_{m+1} \le RA(\gamma|[t_0,t_2^{\tilde e}]) + \tilde b_{m+1}/\sqrt{2}$, so again we have the desired inequality.

Thus, we have shown in all cases that it is possible to construct A such that $RA(A) \le RA(\gamma|[t_0,t_2^{\tilde e}]) + \tilde b_{m+1}/\sqrt{2}$. In the same way, it is possible to choose B such that $RA(B) \le RA(\gamma|[t_3^{\tilde e},t_1]) + \tilde b_{m+1}/\sqrt{2}$.

This finishes the construction of γ_1 connecting y to z such that $RA(\gamma_1) \le RA(\gamma)$, in the case that for each connected component e of $Z_{y,z} \cap C^{(m+1)}$ we have $RA(\gamma|[t_2^e,t_3^e]) \ge \sqrt{2}\,\tilde b_{m+1}$. Thus, we have proved the first conclusion of the Lemma in this case. To finish the proof of the Lemma in this case, we will show that, if $y,z \in [x,\xi]$, then we may choose γ_1 connecting y to z in $Y_{x,\xi}$ so that $RA(\gamma_1) \le RA(\gamma)$ and $\|\gamma_1\|_1 \le (1+\eta^*_m)\|\gamma\|_1$, where $\eta^*_m = O(2^{-m(d-1)/d})$.

In the case when at least one of $RA(\gamma|[t_0,t_2^{\tilde e}])$ or $RA(\gamma|[t_3^{\hat e},t_1])$ is greater than or equal to $\sqrt{2}\,c_{m+1}$, we may use the γ_1 constructed above. When both of these quantities are

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

greater than or equal to $\sqrt{2}\,c_{m+1}$, the γ_1 constructed above is a straight line joining y to z, so clearly $\|\gamma_1\|_1 \le \|\gamma\|_1$.

Now we consider the case when $RA(\gamma|[t_3^{\hat e}, t_1]) \ge \sqrt{2}\,c_{m+1}$ and $RA(\gamma|[t_0, t_2^{\tilde e}]) < \sqrt{2}\,c_{m+1}$. We choose t_7 with $t_5 \le t_7 \le t_4$ such that $\gamma(t_7) \in \mathrm{EDZ}(R^*)$. We let $\tilde\tau$ be the unique point in $\widetilde T$ such that $\pi_1(\tilde\tau) = \pi_1\gamma(t_7)$. Then $\widetilde T = [\tilde\epsilon, \widetilde\omega] = [\tilde\epsilon, \tilde\tau] * [\tilde\tau, \widetilde\omega]$ and the γ_1 that we constructed above in this case satisfies $\gamma_1 = [y, \tilde\epsilon] * \widetilde\gamma_3 * [\tilde\eta, \tilde\tau] * [\tilde\tau, \widetilde\omega] * [\widetilde\omega, \tilde\zeta] * [\tilde\zeta, z]$ or $[y, \tilde\epsilon] * \widetilde\gamma_3 * [\tilde\eta, \tilde\tau] * [\tilde\tau, \widetilde\omega] * \widetilde\gamma_4 * [\tilde\zeta, z]$. It follows from our construction and the fact that $\gamma(t_7)$ is in $\mathrm{EDZ}(r^*)$ that $\|[y, \tilde\epsilon] * \widetilde\gamma_3 * [\tilde\epsilon, \tilde\tau]\|_1 \le (1 + O(2^{-m(d-1)/d}))\|\gamma|[t_0, t_7]\|_1$ and $\|[\tilde\tau, \widetilde\omega] * [\widetilde\omega, \tilde\zeta] * [\tilde\zeta, z]\|_1$, $\|[\tilde\tau, \widetilde\omega] * \widetilde\gamma_4 * [\tilde\zeta, z]\|_1 \le (1 + O(2^{-m(d-1)/d}))\|\gamma|[t_7, t_1]\|_1$. Hence $\|\gamma_1\|_1 \le (1 + O(2^{-m(d-1)/d})\|\gamma\|_1$, as desired.

The case when $RA(\gamma|[t_0, t_2^{\tilde e}]) \ge \sqrt{2}\,c_{m+1}$ and $RA(\gamma|[t_3^{\hat e}, t_1]) < \sqrt{2}\,c_{m+1}$ may be dealt with in the same way.

In the case when $RA(\gamma|[t_0, t_2^{\tilde e}]) < \sqrt{2}\,c_{m+1}$ and $RA(\gamma|[t_3^{\hat e}, t_1]) < \sqrt{2}\,c_{m+1}$, we need to modify the γ_1 constructed above. We recall that $\gamma_1 = A * [\tilde\zeta, \hat\zeta] * B$. Moreover, $A = [y, \tilde\epsilon] * \widetilde\gamma_3 * \widetilde T * [\widetilde\omega, \tilde\zeta]$ or $A = [y, \tilde\epsilon] * \widetilde\gamma_3 * \widetilde T * \widetilde\gamma_4$ and B has a similar description. We let $A_0 = [y, \tilde\epsilon] * \widetilde\gamma_3 * \widetilde T$ and let B_0 be the corresponding modification of B. Thus, A_0 connects y to $\widetilde\omega$ and B_0 connects $\widehat\omega$ to z. We let $\widetilde\gamma_1 = A_0 * [\widetilde\omega, \widehat\omega] * B_0$.

The argument given above shows that $RA(A_0) \le RA(\gamma|[t_0, t_2^{\tilde e}])$. Similarly, $RA(B_0) \le RA(\gamma|[t_3^{\hat e}, t_1])$. Moreover, $RA([\widetilde\omega, \widehat\omega]) = \sqrt{2}\,\tilde b_{m+1} \times$ the number of connected components e of $[x, \xi] \cap C^{(m+1)}$ in $Z_{y,z}$. Thus, $RA(\widetilde\gamma_1) \le RA(\gamma_1)$. The arguments given previously show that $\|\widetilde\gamma_1\|_1 \le 1 + O(2^{-m(d-1)/d}))\|\gamma\|_1$.

In the last construction, we have $\widetilde\gamma_1 = A_0 * [\widetilde\omega, \widehat\omega] * B_0 = [y, \tilde\epsilon] * \widetilde\gamma_3 * \widetilde T * [\widetilde\omega, \widehat\omega] * \widehat T * \widetilde\gamma_3 * [\hat e, z]$. If we set $\widehat\gamma_1 = [y, \hat\zeta] * S\widetilde\gamma_3 * \widehat T * \widehat\gamma_3 * [\widehat\epsilon, z]$, where S is the translation in the direction of $[x, \xi]$ that translates $\widetilde R$ onto $\widehat R$, then $RA(\widehat\gamma_1) \le RA(\widetilde\gamma_1) \le RA(\gamma_1)$, $\|\widehat\gamma_1\|_1 = \|\widetilde\gamma_1\|_1$, and $\widehat\gamma_1$ is a curve in $[x, \xi] \cup \widehat R^*$. Similarly, we may find a curve γ_1^+ in connecting y to z in $[x, \xi] \cup \widetilde R^*$ with $RA(\gamma_1^+) \le RA(\gamma_1)$ and $\|\gamma_1^+\|_1 \le (1 + O(2^{-m(d-1)/d})\|\gamma_1\|_1$.

This finishes the proof of the Lemma in the special case when $RA(\gamma|[t_2^e, t_3^e]) \ge \sqrt{2}\,\tilde b_{m+1}$ for each connected component e of $Z_{y,z} \cap C^{(m+1)}$. We may take γ_1 to be either $\widehat\gamma_1$ or γ_1^+, since both of these curves are in $Y_{\gamma,\xi}$.

Now we suppose that for some connected component e of $[x, \xi] \cap C^{(m+1)}$ that is contained in $Z_{y,z}$ we have $RA(\gamma|[t_2^e, t_3^e]) < \sqrt{2}\,\tilde b_{m+1}$. In this case, there exists $t \in [t_2^e, t_3^e]$ such that $\mathrm{dist}_\infty(\gamma(t), [x, \xi]) \le 3 \times 2^{-n_{m+1}-\sigma-1}$. We choose $t_4 \in [t_2^e, t_3^e]$ to minimize this distance. We let $\pi_i : r_0^\star \to \pi_1^{-1}(\pi_1\gamma(t_i^e))$ be the orthogonal projection, for $i = 2, 3$. We choose $\hat t_2, \hat t_3$ so that $t_2^e < \hat t_2 < \hat t_3 < t_3^e$. We let $\gamma^* : [t_2^e, t_3^e] \to \mathbb{T}^d$ be a curve such that $\gamma^*|[t_2^e, \hat t_2]$ is a reparameterization of $\pi_2\gamma|[t_2^e, t_4]$, $\gamma^*|[\hat t_3, t_3^e]$ is a reparameterization of $\pi_3\gamma|[t_4, t_3^e]$, and $\gamma^*[\hat t_2, \hat t_3]$ is the concatenation of the straight line segment joining $\gamma^*(\hat t_2)$ to $\pi_1\gamma(t_2^e)$, the straight line segment joining $\pi_1\gamma(t_2^e)$ to $\pi_1\gamma(t_3^e)$, and the straight line segment joining $\pi_1\gamma(t_3^e)$ to $\gamma^*(\hat t_3)$. Using the assumption that $RA(\gamma|[t_2, t_3]) < \sqrt{2}\,\tilde b_{m+1}$, it is easy to see that $RA(\gamma^*) \le RA(|[t_2^e, t_3^e])$. Note that γ^* is in $Y_{x,\xi}$, $\gamma^*(t_2^e) = \gamma(t_2^e)$, and $\gamma^*(t_3^e) = \gamma(t_3^e)$.

To prove the existence of γ_1 in general we proceed by induction on the number $n(\gamma)$ of connected components e of $[x, \xi] \cap C^{(m+1)}$ satisfying $e \subset Z_{y,z}$ and $RA(\gamma|[t_2^e, t_3^e]) < \sqrt{2}\,\tilde b_{m+1}$. By what we did above, we have the existence of γ_1 in the case when $n(\gamma) = 0$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

When $n(\gamma) > 0$, we choose a connected component e of $[x, \xi] \cup C^{(m+1)}$ satisfying $e \subset Z_{y,z}$ and $RA(\gamma|[t_2^e, t_3^e]) < \sqrt{2}\,\tilde{b}_{m+1}$. We just showed that there is a curve $\gamma_e^* = \gamma^* : [t_2^e, t_3^e] \to Y_{x,\xi}$ such that $\gamma_e^*(t_2^e) = \gamma(t_2^e)$, $\gamma_e^*(t_3^e) = \gamma(t_3^e)$, and $RA(\gamma^*) \leq RA(\gamma|[t_2^e, t_3^e])$. Our construction above shows that e is part of this curve. Thus we may find $t_4^e \in [t_2^e, t_3^e]$ such that $\gamma_e^*(t_4^e) \in e$. Let $\widetilde{\gamma}^e = \gamma|[t_0, t_2^e] * \gamma_e^*|[t_2^e, t_4^e]$ and $\widehat{\gamma}^e = \gamma_e^*|[t_4^e, t_3^e] * \gamma|[t_4^e, t_1]$.

Both $\widetilde{\gamma}^e$ and $\widehat{\gamma}^e$ are curves satisfying the hypotheses of the Lemma. Moreover, $n(\widetilde{\gamma}^e) + n(\widehat{\gamma}^e) < n(\gamma)$. Hence, by the induction hypothesis, there exist rectifiable curves $\widetilde{\gamma}_1^e : [t_0, t_4^e] \to Y_{x,\xi}$ and $\widehat{\gamma}_1^e : [t_4^e, t_1] \to Y_{x,\xi}$ such that $\widetilde{\gamma}_1^e(t_0) = \widetilde{\gamma}^e(t_0)$; $\widetilde{\gamma}_1^e(t_4^e) = \widetilde{\gamma}^e(t_4^e)$, $RA(\widetilde{\gamma}_1^e) \leq RA(\widetilde{\gamma}^e)$, $\widehat{\gamma}_1^e(t_4^e) = \widehat{\gamma}^e(t_4^e)$, $\widehat{\gamma}_1^e(t_1) = \widehat{\gamma}^e(t_1)$, and $RA(\widehat{\gamma}_1^e) \leq RA(\widehat{\gamma}^e)$. It follows that $\gamma_1 = \widetilde{\gamma}_1^e * \widehat{\gamma}_1^e$ has the required properties.

This proves the first conclusion of the Lemma. To prove the remaining conclusion, it is enough to prove that the curve γ_1 that we have constructed satisfies $\|\gamma_1\|_1 \leq 1 + O(2^{-m(d-1)/d})\|\gamma\|_1$. In fact, this follows easily from what we already did in the case when $n(\gamma) = 0$. We leave the details to the reader. □

12. *Proof of the Proposition in §7: fifth step*

Throughout this section, we suppose that $c_m \geq \nu_m^* b_{m+1}$, for all m. We will prove that in this case $\rho_{L_0,0}(x, y) = \rho^*(x, y)$, for any $x, y \in Z$.

We proved in §7 that $\rho_{L_0,0}(x, y) \leq \rho^*(x, y)$, so it will be enough to prove the opposite inequality here. By the Proposition in §6, $\rho_{L_0,0}(x, y)$ is the infimum of $2^{3/2} \int_\gamma \sqrt{U}\, ds$ over rectifiable curves γ in $\mathbb{T}^d$ connecting x to y. Thus, it will be enough to show that if γ is a rectifiable curve in $\mathbb{T}^d$ connecting x to y, then $\rho^*(x, y) \leq 2^{3/2} \int_\gamma \sqrt{U}\, ds$.

By the Lemma in §8, we may assume, without loss of generality, that γ is in J^*.

If there is no tooth R of C such that $x, y \in R^*$, we set $m_0 = -1$. Otherwise, we choose m_0 as large as possible so that there is an ultimate tooth $R = R^{(m_0)}$ of $C^{(m_0)}$ with $x, y \in R^{(m_0)^*}$. In this case, we have $m_0 \geq 0$. Note that $Z \cap R^* = C_R^{(\infty)}$, so the hypothesis that $x, y \in R^*$ is equivalent to $x, y \in C_R^{(\infty)}$, under our standing assumption that $x, y \in Z$.

In the case when $m_0 \geq 0$, there exists a rectifiable curve γ_1 in $R^{(m_0)*}$ such that $RA(\gamma_1) \leq \sqrt{2}\,\nu_{m_0}^* b_{m_0+1} \leq \sqrt{2}\,c_{m_0} < \sqrt{2}\,b_{m_0}$. Consequently, it follows from the Lemma in §9 that there exists a rectifiable curve γ_1 in $R^{(m_0)*}$ such that $RA(\gamma_1) \leq RA(\gamma)$. Thus, we may assume, without loss of generality, that γ is in $R^{(m_0)*}$.

We let $Z_{x,y}^*$ be as in §7, i.e. it is the shortest Jordan arc in $\widehat{Z}_{x,y}$ joining x to y. In §7, we showed that $Z_{x,y}^*$ can be decomposed as $Z_{x,\xi}^* \cup [\xi, \eta] \cup Z_{\eta,y}^*$, where ξ and η are defined differently in different cases. When $m_0 = -1$, $[\xi, \eta]$ is a line segment in J. If, in addition, $x \in J \cap C$ (respectively $y \in J \cap C$), then $x = \xi$ (respectively $y = \eta$). When $m_0 \geq 0$, $[\xi, \eta]$ is a sub-arc of $\partial R^{(m_0)}$. If, in addition, $x \in C^{(m_0)} \cap C^{(m_0+1)}$ (respectively $y \in C^{(m_0)} \cap C^{(m_0+1)}$), then $x = \xi$ (respectively $y = \eta$).

If ξ and η are in the same side of $R^{(m_0)}$, or both are in the top of $R^{(m_0)}$, then $[\xi, \eta]$ is a line segment in the relevant side or in the top. If one of ξ or η is in a side of $R^{(m_0)}$ and the other is in the top, then $[\xi, \eta] = [\xi, \omega] \cup [\omega, \eta]$, where $[\xi, \omega]$ and $[\omega, \eta]$ are line segments and ω is the top of a side of $R^{(m_0)}$. If ξ and η are in different sides of $R^{(m_0)}$, then $[\xi, \eta] = [\xi, \omega] \cup [\omega, \nu] \cup [\nu, \eta]$, where $[\xi, \omega]$, $[\omega, \nu]$, and $[\nu, \eta]$ are line segments, ω is the top of the side that contains ξ, and ν is the top of the side that contains η.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

In the case when ξ and η are on different sides of $R^{(m_0)}$, there exist rectifiable curves γ_1 and γ_2 in $R^{(m_0)^*}$ such that γ_1 connects x to ω, γ_2 connects ν to y, and

$$RA(\gamma_1 * [\omega, \nu] * \gamma_2) \leq RA(\gamma).$$

This follows from the arguments in §10, together with the assumption that $c_{m_0} \geq \nu^*_{m_0} b_{m_0+1}$. Thus, we may assume, without loss of generality, that $\gamma = \gamma_1 * [\omega, \nu] * \gamma_2$. Similarly, if ξ is in a side of $R^{(m_0)}$ and η is in the top of $R^{(m_0)}$, there exists a rectifiable curve γ_1 connecting x to ω in $R^{(m_0)^*}$ such that

$$RA(\gamma_1 * [\omega, \eta]) \leq RA(\gamma).$$

We may assume, without loss of generality, that $\gamma = \gamma_1 * [\omega, \eta]$.

Next, we consider the case when ξ and η are both in the same side of $R^{(m_0)}$. We let u denote the bottom endpoint of this side and ω the top endpoint. Then $x, y \in Z_{u,\omega}$ and the Lemma of §11 applies to this situation. It follows that there exists a rectifiable curve γ_1 in $Y_{u,\omega}$ connecting x to y such that $RA(\gamma_1) \leq RA(\gamma)$. Thus, we may assume, without loss of generality, that γ is in $Y_{u,\omega}$.

We may apply the Lemma in §11 in the same way to the other cases. In the case that ξ and η are in different sides of $R^{(m_0)}$, we let u (respectively v) denote the bottom endpoint and ω (respectively ν) denote the top endpoint of the side that contains ξ (respectively η). We showed above that we may assume that $\gamma = \gamma_1 * [\omega, \nu] * \gamma_2$, where γ_1 connects x to ω and γ_2 connects ν to y. By the Lemma in §11, we may assume, without loss of generality, that γ_1 is in $Y_{u,\omega}$ and γ_2 is in $Y_{v,v}$. Similarly, in the case that ξ is in a side of $R^{(m_0)}$ and η is in the top, we may assume that $\gamma = \gamma_1 * [\omega, \eta]$ where γ_1 is in $Y_{u,\omega}$.

We may also apply the Lemma in §11 when $m_0 = -1$. Let u and ω be the endpoints of J. In the case that $m_0 = -1$, it follows from the Lemma in §11 that we may assume that γ is in $Y_{u,\omega}$.

The proof of the Lemma in §11 gives more information concerning the curve (γ, γ_1, or γ_2, according to the case) in $Y_{u,\omega}$ (or $Y_{v,v}$). It has the form $\mu_{j_0} * \mu_{j_0+1} * \cdots * \mu_{j_1}$, where $j_0 = 0$ or 1 and $j_1 = 2k$ or $2k+1$, and k is an integer greater than or equal to 1. For odd i, μ_i is a curve in R_i^*, where R_i is an ultimate tooth of $C^{(m_0+1)}$ whose bottom is in $[u, \omega]$ (or $[v, v]$). For even i, μ_i is a line segment in $[u, \omega]$ (or $[v, v]$). For odd i, $i \neq 1, 2k+1$, $\mu_i = \mu_i^1 * t_i * \mu_i^2$, where μ_i^1 connects the bottom to the top of one side of R_i, μ_i^2 connects the top to the bottom of the other side of R_i, and t_i is the top of R_i. The curve μ_i also has this form in the case when $i = 1$ and $j_0 = 0$.

Let $\Omega^m_{x,y}$ denote the union of $\widehat{Z}_{x,y}$ (defined in §7) and all R^* where R is an ultimate tooth of $C^{(m+1)}$ such that $R^* \cap \widehat{Z}_{x,y}$ contains more than one point.

In the case when $m_0 \geq 0$, we have $x, y \in R^{(m_0)^*}$. In this case, $\Omega^{(m_0-1)}_{x,y} = R^{(m_0)^*}$.

In the case when $m_0 = -1$, we have $x, y \in J \setminus C$. In this case $\Omega^{m_0}_{x,y}$ is the union of the line segment $[x, y]$ and all R^* such that R is a tooth of C whose bottom is in $[x, y]$.

In the case when $m_0 \geq 0$, we described above a decomposition $Z^*_{x,\xi} \cup [\xi, \eta] \cup Z^*_{\eta,y}$ of $Z^*_{x,y}$, where possibly $x = \xi$ and possibly $y = \eta$. If $x \neq \xi$, we let $R_x^{(m_0+1)}$ denote the ultimate tooth of $C^{(m_0+1)}$ such that $x \in R_x^{(m_0+1)}$. Likewise, if $y \neq \eta$, we let $R_y^{(m_0+1)}$ denote the ultimate tooth of $C^{(m_0+1)}$ such that $y \in R_y^{(m_0+1)^*}$. In the case when $m_0 \geq 0$,

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

$\Omega^{m_0}_{x,y}$ consists of $[\xi, \eta]$, all R^* such that R is an ultimate tooth of $C^{(m_0+1)}$ whose bottom is in $[\xi, \eta]$, $R^{(m_0+1)^*}_x$ if $x \neq \xi$, and $R^{(m_0+1)^*}_y$ if $y \neq \eta$.

The argument given above shows that there exists a rectifiable curve γ_1 in $\Omega^{m_0}_{x,y}$ connecting x to y such that $RA(\gamma_1) \leq RA(\gamma)$. Thus, for proving the Proposition in §7, we may assume, without loss of generality, that γ is in $\Omega^{m_0}_{x,y}$.

In showing this, we decomposed γ, γ_1, or γ_2 (according to the case we are considering) as $\mu_{j_0} * \mu_{j_1} * \cdots * \mu_{j_1}$. For odd i, R_i is a curve in R^*_i, where R_i is an ultimate tooth of $C^{(m_0+1)}$ whose bottom is in $[\xi, \eta]$. We may apply the argument we gave above for μ_i (i odd) in place of γ and R^*_i in place of $R^{(m_0)^*}$. In this way, we may replace μ_i (i odd) with a curve in $R^*_i \cap \Omega^{(m_0+1)}_{x,y}$ having the same endpoints as μ_i and having reduced action less than or equal to the reduced action for μ_i.

Doing this for each odd i, we obtain a rectifiable curve γ_1 in $\Omega^{(m_0+1)}_{x,y}$ having the same endpoints as γ with $RA(\gamma_1) \leq RA(\gamma)$.

Repeating this construction enough times, we obtain a rectifiable curve γ_p in $\Omega^p_{x,y}$ having the same endpoints as γ and satisfying $RA(\gamma_p) \leq RA(\gamma)$, for any $p \geq m_0$.

Note that $\Omega^p_{x,y} \supset \Omega^{p+1}_{x,y}$, for all p, and $\bigcap_p \Omega^p_{x,y} = \widehat{Z}_{x,y}$. In the construction above, γ_{p+1} is obtained by altering γ_p. The altered parts of γ_p are each in some R^*, where R is an ultimate tooth of $C^{(p+1)}$. Consequently, the sequence $\ldots, \gamma_p, \ldots$ converges to a Jordan arc γ_∞, with respect to the Hausdorff metric. (Here, we identify γ_p with its image in $\mathbb{T}^d$.)

We claim that it is possible to construct the γ_p's so that not only does each γ_p have finite arc-length and $RA(\gamma_{p+1}) \leq RA(\gamma_p)$, but also the arc-lengths of the γ_p's are bounded above. We observed above that γ_{p+1} is obtained by altering various parts of γ_p. Each such part μ is in an R^*, where R is an ultimate tooth of $C^{(p+1)}$. When p is large enough so that $x, y \in C^{(p)} \smallsetminus C^{(p+1)}$, such a part μ of γ_p has the form $\mu = \mu_1 * T * \mu_2$, where μ_1 connects the bottom endpoint of one side of R to the top endpoint, T is the top of R, and μ_2 connects the top endpoint of the other side of R to the bottom endpoint. The replacement $\tilde{\mu}$, given by the Lemmas in §10 and §11, satisfies $\|\tilde{\mu}\|_1 \leq (1 + O(2^{-p(d-1)/d})\|\mu\|_1$. Because this is true for each replacement part, it is also true for the whole, i.e. $\|\gamma_{p+1}\|_1 \leq (1 + O(2^{-p(d-1)/d})\|\gamma_p\|$. Since this is true for all sufficiently large p, it follows that there is an upper bound on $\|\gamma_p\|_1$, as asserted.

To summarize, we have proved the following.

LEMMA. *Let $x, y \in Z$ and suppose that $c_m \geq \nu^*_m b_{m+1}$ for all m. Let γ be a rectifiable curve in $\mathbb{T}^d$ connecting x to y. Then there exists a rectifiable Jordan curve γ_1 in $\widehat{Z}_{x,y}$ connecting x to y such that $RA(\gamma_1) \leq RA(\gamma)$.* □

Proof that $\rho_{L_0,0} = \rho^*$ in the case when $c_m \geq \nu^*_m b_{m+1}$ for all m. We proved in §7 that $\rho_{L_0,0} \leq \rho^*$. Thus, it is enough to consider a rectifiable curve γ in $\mathbb{T}$ connecting x to y and prove that $RA(\gamma) \geq \rho^*(x, y)/2$. In view of the Lemma above, we may suppose that γ is a rectifiable Jordan curve in $\widehat{Z}_{x,y}$.

Since γ is a Jordan arc in $\widehat{Z}_{x,y}$, we have a decomposition $\gamma = \mu_{j_0} * \mu_{j_0+1} * \cdots * \mu_{j_1}$, where $j_0 = 0$ or 1, $j_1 = 2k$ or $2k + 1$, k is an integer greater than or equal to 0, μ_i is a connected component of $Z_{x,y} \cap Z^*_{x,y}$ for even i, and the endpoints of μ_i are the endpoints of a primary gap G_i of $Z^*_{x,y}$ for odd i. Clearly, $RA(\mu_i) = 0$ for even i.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

We claim that $RA(\mu_i) \geq \sqrt{2}\, b^{(\infty)}_{m(G_i)}$ for odd i. In view of the definition of $\rho^*(x, y)$, this implies $RA(\gamma) \geq \rho^*(x, y)/2$.

Let ξ and η be the endpoints of G_i. We have that μ_i is a Jordan arc in $\widehat{Z}_{\xi,\eta}$ connecting ξ and η. Let β_p be the number of bottoms of ultimate teeth in $C^{(p+1)}$ that are in μ_i. We have $\beta_p = 0$ for $p < m(G_i)$. If $\mu_i = G_i$, we have $\beta_{m(G_i)} = 1$ and $\beta_p = 0$ for all $p \neq m(G_i)$. We let $\delta_m = 1$, and

$$\delta_p = \nu^*_m \cdots \nu^*_{p-1}\beta_m + \nu^*_{m+1} \cdots \nu^*_{p-1}\beta_{m+1} + \cdots + \nu^*_{p-1}\beta_{p-1} + \beta_p,$$

for $p > m$, where $m = m(G_i)$. If $m \leq q < p$, we have that $\nu^*_q \cdots \nu^*_{p-1}$ is the number of ultimate teeth of $C^{(p+1)}$ that lie in R^*, where R is an ultimate tooth of $C^{(q+1)}$. Consequently, $\delta_p \leq \nu^*_m \cdots \nu^*_{p-1}$. Moreover $RA(\mu_i) \geq \sqrt{2}\,(\beta_m c_m + \beta_{m+1}c_{m+1} + \cdots + \beta_{p-1}c_{p-1} + \beta_p c_p) \geq \sqrt{2}\,\delta_p \nu^*_p b_{p+1}$, in view of the assumption that we made in this section that $c_n \geq \nu^*_n b_{n+1}$, for all n.

The length of the bottom of an ultimate tooth of $C^{(p+1)}$ is 2^{-n_p}. We have $\nu^*_q = 2^{6+q} - 2$. Thus $\nu^*_m \cdots \nu^*_{p-1} 2^{-n_p} \to \infty$ as $p \to \infty$. The total length of all the bottoms of ultimate teeth of $C^{(p+1)}$ is $\nu^*_m \cdots \nu^*_{p-1} 2^{-n_p}$. Among such teeth r we consider those whose top is contained in μ_i. The number of such teeth is $\nu^*_m \cdots \nu^*_{p-1} - \delta_p$. The total length of these is $(\nu^*_m \cdots \nu^*_{p-1} - \delta_p)2^{-n_p}$ and it is bounded by the arc-length of μ_i. Since $(\nu^*_m \cdots \nu^*_{p-1} - \delta_p)2^{-n_p}$ is bounded independently of p and $\nu^*_m \cdots \nu^*_{p-1} 2^{-n_p} \to \infty$ as $p \to \infty$, it follows that $(\nu^*_m \cdots \nu^*_{p-1} - \delta_p)/\nu^*_m \cdots \nu^*_{p-1} \to 0$ as $p \to \infty$, i.e. $\delta_p/\nu^*_m \cdots \nu^*_{p-1} \to 1$ as $p \to \infty$. Since $RA(\mu_i) \geq \sqrt{2}\,\delta_p \nu^*_p b_{p+1}$ and $b_m, \nu^*_m b_{m+1}, \nu^*_m \nu^*_{m+1} b_{m+2} \cdots$ is a decreasing sequence (in the case that $c_n \geq \nu^*_n b_{n+1}$ for all n), it follows that $RA(\mu_i) \geq \sqrt{2}\, b^{(\infty)}_m$. As we observed above, this implies $RA(\gamma) \geq \rho^*(x, y)/2$. □

13. *End of the proof of the Proposition in §7*

Proof. In §7, we showed that $\rho_{L_0,0} \leq \rho^*$.

In §12, we showed that $\rho_{L_0,0} = \rho^*$ in the case that $c_m \geq \nu^*_m b_{m+1}$ for all m. To prove the Proposition in §7, all that remains is to show that, in general, we have $\rho_{L_0,0}(x, y) > 0$ for any $x, y \in Z$ such that $\rho^*(x, y) > 0$.

In §5, we defined $c^{(\infty)}_m = \inf(c_m, \nu^*_m c_{m+1}, \nu^*_m \nu^*_{m+1} c_{m+2}, \ldots)$. Clearly $c_m > 0$ and $c^{(\infty)}_m = \min(c_m, \nu^*_m c^{(\infty)}_{m+1})$. Hence $c^{(\infty)}_m = 0$ if and only if $c^{(\infty)}_{m+1} = 0$. Moreover, it follows from the definitions of b_m and c_m that $c_m/b_m \to 1$, as $m \to \infty$. Hence $c^{(\infty)}_m > 0$ (for one, and therefore for every, m) if and only if $b^{(\infty)}_m > 0$ (for one, and therefore for every, m).

Consider $x, y \in Z$ such that $\rho^*(x, y) > 0$. The existence of such x, y implies that $c^{(\infty)}_m > 0$ for all m. Since $\rho^*(x, y) > 0$, there exists a primary gap G of $Z^*_{x,y}$. Let γ be a rectifiable curve in $\mathbb{T}^d$ connecting x to y. We will show that $RA(\gamma) \geq \sqrt{2}\, c^{(\infty)}_{m(G)}$. This will be enough to finish the proof, since it implies that $\rho_{L_0,0}(x, y) \geq \sqrt{2}\, c^{(\infty)}_{m(G)} > 0$.

To show that $RA(\gamma) \geq \sqrt{2}\, c^{(\infty)}_{m(G)}$, we will suppose the contrary and obtain a contradiction.

If there is no tooth R of C such that $x, y \in R^*$, we set $m_0 = -1$. Otherwise, we choose m_0 as large as possible so that there is an ultimate tooth $R^{(m_0)}$ of $C^{(m_0)}$ with $x, y \in R^{(m_0)^*}$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

By the Lemma in §8, we may assume that γ is in J^*. Suppose $m_0 \geq 0$. By the Lemma in §9, there exists a rectifiable curve γ_1 in $R^{(m_0)^*}$ such that $RA(\gamma_1) \leq RA(\gamma)$, since $RA(\gamma) < \sqrt{2}\, c^{(\infty)}_{m(G)} < \sqrt{2}\, b_{m_0}$. Hence, we may assume that γ is in $R^{(m_0)^*}$, in the case when $m_0 \geq 0$.

Now we proceed as in §12. Thus, we have $Z^*_{x,y} = Z^*_{x,\xi} \cup [\xi, \eta] \cup Z^*_{\eta,y}$, with the same discussion as in §12. In the case when ξ and η are on different sides of $R^{(m_0)}$, we again have the existence of curves γ_1 and γ_2 with properties stated in §12. As in §12, the existence of these curves follows from the arguments in §10, but with the assumption that $RA(\gamma) < \sqrt{2}\, c^{(\infty)}_{m(G)} \leq \sqrt{2}\, c_{m_0}$ used in place of the assumption that $c_{m_0} \geq \nu^*_{m_0} b_{m_0+1}$.

In the case when ξ is in a side of $R^{(m_0)}$ and η is in the top of $R^{(m_0)}$, we have the existence of a curve γ_1 having the properties listed in §12.

Next, we consider the case when ξ and η are both in the same side of $R^{(m_0)}$. We let u denote the bottom endpoint of this side and ω the top endpoint. Then $x, y \in Z_{u,\omega}$ and $\xi, \eta \in [u, \omega]$.

A slight modification of the proof of the Lemma in §11 proves that there exists a rectifiable curve γ_1 in $Y_{u,\omega}$ connecting x to y such that $RA(\gamma_1) \leq RA(\gamma)$. The modified argument uses the assumption that $RA(\gamma) < \sqrt{2}\, c^{(\infty)}_{m(G)}$. In the notation of §11, we have

$$RA(\gamma|[t^e_2, t^e_3]) \leq RA(\gamma) < \sqrt{2}\, c^{(\infty)}_{m(G)} \leq \sqrt{2}\, c_{m_0+1} < \sqrt{2}\, \tilde{b}_{m_0+1},$$

for each connected component e of $[u, \omega] \cap C^{(m_0+1)}$ that is contained in $Z_{x,y}$. As in §11, there exists $t \in [t^e_2, t^e_3]$ such that $\mathrm{dist}_\infty(\gamma(t), [u, \omega]) \leq 2^{-n_{m_0+1}-\sigma}$. As in §11, it follows that we may replace $\gamma|[t^e_2, t^e_3]$ by $\gamma^*_e : [t^e_2, t^3_3] \to Y_{u,\omega}$ with $\gamma^*_e(t^e_2) = \gamma(t^e_2), \gamma^*_e(t^e_3) = \gamma(t^e_3)$, and $RA(\gamma^*_e) \leq RA(\gamma|[t^e_2, t^e_3])$. Let γ_1 be the curve obtained by replacing each $\gamma|[t^e_2, t^e_e]$ with γ^*_e. Clearly, γ_1 is a curve in $Y_{u,\omega}$ connecting x to y and $RA(\gamma_1) \leq RA(\gamma)$. Hence, we may assume that γ is in $Y_{u,\omega}$.

As in §12, we may argue the same way in the other cases. In this way, we get the same results as in §12.

If ξ and η are in different sides of $R^{(m_0)}$, then we may assume that $\gamma = \gamma_1 * [\omega, v] * \gamma_2$ with γ_1 in $Y_{u,\omega}$ and γ_2 in $Y_{v,v}$. If ξ is in a side of $R^{(m_0)}$ and η is in the top, we may assume that $\gamma = \gamma_1 * [\omega, \eta]$ where γ_1 is in $Y_{u,\omega}$ (notation as in §12).

Now we suppose that there is at least one ultimate tooth $R^{(m_0+1)}$ of $C^{(m_0+1)}$ whose bottom is in $[\xi, \eta]$. Let δ and ϵ be the endpoints of the bottom of $R^{(m_0+1)}$. The argument above shows that we may assume that $\gamma = \gamma_1 * \gamma_2 * \gamma_3$, where γ_1 connects x to δ, γ_2 connects δ to ϵ in $R^{(m_0+1)}$, and γ_3 connects ϵ to y. Below, we will show that $RA(\gamma_2) \geq \sqrt{2}\, c^{(\infty)}_{m_0+1}$. This will give the desired contradiction to our assumption that $RA(\gamma) < \sqrt{2}\, c^{(\infty)}_{m(G)}$, since $m(G) \geq m_0 + 1$.

By the Lemma in §10, $RA(\gamma_2) \geq \sqrt{2}\, c_{m_0+1} \geq \sqrt{2}\, c^{(\infty)}_{m_0+1}$ or there exists a rectifiable curve $\gamma^*_2 = \alpha * T * \beta$ in $R^{(m_0+1)^*}$, where α connects δ to the top δ_1 of the side of $R^{(m_0+1)}$ that contains δ, T is the top of $R^{(m_0+1)}$, and β connects the top of the other side of $R^{(m_0+1)}$ to ϵ. In the first case we are done, so we may assume the existence of γ^*_2 with the properties listed above.

Now we apply the Lemma in §11 to α. Hypothesis (b) of this Lemma is satisfied by α. By the conclusion, there exists a rectifiable curve α_1 connecting δ to δ_1 with $RA(\alpha_1) \leq RA(\alpha)$ and $\|\alpha_1\|_1 \leq 1 + O(2^{-m_0(d-1)/d})$ in Y_{δ,δ_1}. We may similarly apply

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

the Lemma in §11 to β and get a curve β_1 with the same properties (relative to β rather than to α) as α_1. It follows that we may assume, without loss of generality, that $\gamma_2 = \mu_1 * \cdots * \mu_k$, $k = 2\nu^*_{m_0+1} - 1$, where for i even μ_i is a connected component of $C^{(m_0+2)} \cap \partial R^{(m_0+1)}$ and for i odd μ_i is in ρ_i^* where ρ_i is an ultimate tooth of $C^{(m_0+2)}$ whose bottom is in $\partial R^{(m_0+1)}$.

Now we repeat this process on the next level. If $RA(\mu_i) \geq \sqrt{2}\, c_{m_0+2}$, we leave μ_i as it is. Otherwise, we replace μ_i with a curve passing through the four vertices of ρ_i, using the Lemma in §10. We then modify this new curve using the Lemma in §11, just as in the previous step.

Next we repeat this process infinitely often, each time on one higher level. Each time the reduced action is not increased and the L^1 length is increased by a factor bounded by $1 + O(2^{-p(d-1)/d})$ at the pth level. This process converges to a rectifiable curve of finite arc-length, which we will denote by γ_2^∞. We will also denote a curve at the pth level by γ_2^p. We have $RA(\gamma_2^\infty) \leq RA(\gamma_2^p) \leq RA(\gamma_2)$ and $\|\gamma_2^p\|_1$ is bounded above independently of p.

An argument similar to that at the end of §12 shows that $RA(\gamma_2^\infty) \geq c^{(\infty)}_{m_0+1}$. Since $m(G) \geq m_0 + 1$ and $RA(\gamma) \geq RA(\gamma_2) \geq RA(\gamma_2^\infty)$, this gives the desired contradiction, in the case when there is at least one ultimate tooth of $R^{(m_0+1)}$ of $C^{(m_0+1)}$ whose bottom is in $[\xi, \eta]$. Note that this is the case if $m(G) = m_0 + 1$, because G is the bottom of an ultimate tooth $R^{(m_0+1)}$ of $C^{(m_0+1)}$.

In general, we get the desired result by induction on $m(G) - m_0 - 1$. We have just obtained it when $m(G) - m_0 - 1 = 0$. Since G is in $Z^*_{x,y} = Z^*_{x,\xi} \cup [\xi, \eta] \cup Z^*_{\eta,y}$, we have that G is in $Z^*_{x,\xi}$ or $Z^*_{\eta,y}$ in the case when $m(G) - m_0 - 1 > 0$. We may assume, without loss of generality, that $\gamma = \gamma_1 * [\xi, \eta] * \gamma_2$. Here γ_1 (respectively γ_2) is reduced to the single point x (respectively y) if $x = \xi$ (respectively $y = \eta$). If $x \neq \xi$ (respectively $y \neq \eta$) then γ_1 (respectively γ_2) connects x to ξ (respectively η to y) in $R_x^{(m_0+1)}$ (respectively $R_y^{(m_0+1)^*}$), where $R_x^{(m_0+1)}$ (respectively $R_y^{(m_0+1)}$) is the ultimate tooth of $C^{(m_0+1)}$ that contains x (respectively y).

Now we suppose that G is in $Z^*_{x,\xi}$, the case when G is contained in $Z^*_{\eta,y}$ being similar. In this case $x \neq \xi$. We claim that the inductive hypothesis implies that $RA(\gamma_1) \geq \sqrt{2}\, c^{(\infty)}_{m(G)}$. Indeed the conditions on x and y that we discussed at the beginning of this section are satisfied for x and ξ, but with m_0 replaced by $m_0 + 1$, so we may apply the inductive hypothesis. □

14. *End of the proof of the Proposition in §5*

If $b_m^{(\infty)} = 0$ for some m (and therefore for every m), then ρ^* vanishes identically, by its definition. Since $\rho_{L_0,0} \leq \rho^*$, it follows that $\rho_{L_0,0}$ vanishes identically. Hence $\bar{A}_{L_0,0}$ is reduced to one point.

If $b_m^{(\infty)} \neq 0$ for some m (and therefore for every m), then we may choose $x, y \in Z$ such that $Z^*_{x,y}$ contains at least one primary gap. Then $\rho^*(x, y) > 0$ by the definition of ρ^*. By the Proposition in §7, $\rho_{L_0,0}(x, y) > 0$. Hence $\bar{A}_{L_0,0}$ is homeomorphic to an interval.

In the case when $b_m^{(\infty)} \neq 0$ and $c_m \geq \nu^*_m b_{m+1}$ for all m, then we have $\rho_{L_0,0} = \rho^*$ by the Proposition in §7. Hence, $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is the quotient metric space associated to

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

ρ^* in §7. As we observed in §7, in this case the quotient metric space associated to ρ^* is isometric to an interval, so we have proved that $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval in the case that $b_m^{(\infty)} \neq 0$ and $c_m \geq \nu_m^* b_{m+1}$ for all m.

Suppose that $c_m \geq \nu_m^* b_{m+1}$ for all $m \geq m_0$. Let x and y be the bottom endpoints of the sides of an ultimate tooth R of $C^{(m_0)}$. In this case, the discussion in §12 shows that $\rho_{L_0,0}(\xi, \eta) = \rho^*(\xi, \eta)$ for any $\xi, \eta \in Z_{x,y}$. Hence, the image of $Z_{x,y}$ in $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval. Now suppose, in addition, that $b_{m_0-1} < \nu_{m_0-1}^* b_{m_0}^{(\infty)}$. Let R be an ultimate tooth of $C^{(m_0-1)}$. Let x and ξ be the bottom and top endpoints of one side of R, and y and η the bottom and top endpoints of the other side. Then $\rho_{L_0,0}(x, \xi) = \rho^*(x, \xi) = \rho^*(\xi, y) = \rho_{L_0,0}(\xi, y) = \nu_{m_0-1}^* b_{m_0}^{(\infty)}/2$ and $\rho_{L_0,0}(x, y) < b_{m_0-1}$. Hence $\rho_{L_0,0}(x, y) < b_{m_0-1} < \nu_{m_0-1}^* b_{m_0}^{(\infty)} = \rho_{L_0,0}(x, \xi) + \rho_{L_0,0}(\xi, y)$ and $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is not isometric to an interval.

Finally, if R is an ultimate tooth of some $C^{(m)}$ and x, ξ, η, y are as in the previous paragraph in relation to R, then $\rho_{L_0,0}(x, \xi) \geq \nu_m^* c_{m+1}^{(\infty)}/2 = \rho_{L_0,0}(\xi, y)$ and $\rho_{L_0,0}(x, y) \leq b_m$, so if $b_m < \nu_m^* c_{m+1}^{(\infty)}$, then $\rho_{L_0,0}(x, y) < \rho_{L_0,0}(x, \xi) + \rho_{L_0,0}(\xi, y)$. If $b_m < \nu_m^* c_{m+1}^{(\infty)}$ for infinitely m, any segment of $\bar{A}_{L_0,0}$ contains such pairs x, y, so $\bar{A}_{L_0,0}$ does not contain any sub-set isometric to an interval. □

15. *Proof of Theorem 2*

Theorem 2 follows easily from Theorem 1 by an application of the Maupertuis–Carneiro principle. We begin by discussing the Maupertuis principle and a refinement of it, due to Carneiro.

Let g be a Riemannian metric on a closed manifold M. Let $K = g/2$ denote the kinetic energy associated to g and let $U : M \to \mathbb{R}$ be a potential. The mechanical system associated to U has Lagrangian $L = K + U$ and Hamiltonian $H = K - U$, where we identify the cotangent bundle of M with the tangent bundle as usual by means of Legendre transformation. In this case, the Legendre transformation is the canonical isomorphism associated to g. Since the system is autonomous, H is preserved by the Lagrangian flow associated to L (which is the same as the Hamiltonian flow associated to H). Thus, if $E \in \mathbb{R}$, the energy hypersurface $\{H = E\}$ is preserved by the flow.

The Maupertuis principle states that the Euler–Lagrange trajectories of L on the energy hypersurface $\{H = E\}$ are the same, except for a time reparameterization, as the geodesics of the Riemannian metric $(E + U)g$. (This is a Riemannian metric on the open sub-set $\{U < E\}$ of M.) The Lagrangian for the geodesic flow of the metric $(E + U)g$ is $(E + U)K$ and the Hamiltonian is $K/(E + U)$. Here, again, we use the identification of the cotangent bundle of $\{U < E\}$ with the tangent bundle given by the Legendre transformation. In this case, the Legendre transformation is the isomorphism canonically associated to the Riemannian metric $(E + U)g$ on $\{U < E\}$.

The infinitesimal generator of the geodesic flow associated to the metric $(E + U)g$ is the symplectic gradient of $K/(E + U)$, i.e. the vector field corresponding to the 1-form $d(K/(E + U)) = ((E + U)\,dK - K\,dU)/(E + U)^2$ under the canonical correspondence between vector fields and 1-forms on $T^*\{U < E\}$ associated to the canonical symplectic form on $T^*\{U < E\}$. On the energy hypersurface $\{H = E\}$, this

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

1-form is $(dK - dU)/K = dL/K$. It follows that the geodesic vector field associated to the metric $(E + U)g$ is $1/K\times$ the Euler–Lagrange vector field associated to L on the energy hypersurface $\{H = E\}$. Thus, if $\gamma(t)$ is a Euler–Lagrange trajectory of L on the energy hypersurface $\{H + E\}$, then $\gamma(\tau)$ is a geodesic of $(E + U)g$, where $d\tau/dt = K$.

Carneiro [**C**] showed that c-minimizers of L lie on the energy level $E = \alpha(c)$ and that, under the Maupertuis principle, they correspond to c-minimizers of $K^* = (\alpha(c) + U)K$.

In fact, let $\gamma : [t_0, t_1] \to M$ be an absolutely continuous curve of constant energy $\alpha(c)$, i.e. such that $K(\gamma(t), \dot\gamma(t)) - U(\gamma(t)) \equiv \alpha(c)$. Let η be a closed 1-form on M whose deRham cohomology class is c. Then

$$\begin{aligned}\int_{t_0}^{t_1} (K + U - \eta + \alpha(c)) \left(\gamma(t), \frac{d\gamma}{dt}(t)\right) dt \\ = \int_{\tau_0}^{\tau_1} \left(K\frac{d\tau}{dt} + (U + \alpha(c))\frac{dt}{d\tau} - \eta\right) \left(\gamma(\tau), \frac{d\gamma}{d\tau}(\tau)\right) d\tau \\ = \int_{\tau_0}^{\tau_1} (K^* + 1 - \eta) \left(\gamma(\tau), \frac{d\gamma}{d\tau}(\tau)\right) d\tau\end{aligned}$$

if $d\tau/dt = U + \alpha(c)$. In this way, (regular) c-minimizers correspond to (regular) c-minimizers for K^*. This calculation also shows that $\rho_{K^*,c} = \rho_{L,c}$ (cf. [**C**]).

Let $U : \mathbb{T}^d \to \mathbb{R}$ be a non-negative, somewhere vanishing function of class $C^{2d-3,1-\epsilon}$ on $\mathbb{T}^d$ for which the conclusion of Theorem 1 holds. Let $\pi : \mathbb{T}^{d+1} \to \mathbb{T}^d$ be defined by $\pi(\theta_0, \theta_1, \dots, \theta_d) = (\theta_1, \dots, \theta_d)$. Let

$$\tilde L_0 = \frac{1}{2}\sum_{i=0}^{n} d\theta_i^2 + U \circ \pi.$$

Let $c = [d\theta_0] \in H^1(\mathbb{T}^{d+1}; \mathbb{R})$. Let $Z \subset \mathbb{T}^d$ be the zero set of U. The regular c-minimizers of $\tilde L_0$ satisfy $d\theta_0/dt = 1$ and $(\theta_1(t), \dots, \theta_n(t)) = \text{constant} \in Z$. Thus, $A_{\tilde L_0,c} = \mathbb{T} \times A_{L_0,0} = \mathbb{T} \times Z \subset \mathbb{T}^{d+1}$. It is easily seen that $\rho_{\tilde L_0,c} = \rho_{L_0,0} \circ \pi$. The c-minimizers have energy $E = 1/2$, so $\alpha(c) = 1/2$.

Let $g = (1/2 + U)\sum_{i=0}^{d} d\theta_i^2$ and let $K = g/2$ be the kinetic energy associated to g. By Carneiro's result, $\rho_{K,c} = \rho_{\tilde L_0,c} = \rho_{L_0,0} \circ \pi$. Hence g is a Riemannian metric of class $C^{2d-3,1-\epsilon} = C^{2(d+1)-5,1-\epsilon}$ on $\mathbb{T}^{d+1}$ such that $(\bar A_{K,c}, \rho_{K,c})$ is isometric to $(\bar A_{L_0,0}, \rho_{L_0,0})$. If U is chosen so that the conclusion of Theorem 1 holds, then $(\bar A_{K,c}, \rho_{K,c})$ is isometric to an interval and we have the conclusion of Theorem 2. □

16. *Some connecting orbits*

Throughout this section, we suppose that L is a Lagrangian in the sense defined in §1. This means that L is a C^2 real-valued function defined on $TM \times \mathbb{T}/\mathbb{Z}$ that satisfies a Legendre positive definiteness condition, superlinear growth in the fiber, and completeness of the Euler–Lagrange flow. We suppose $c \in H^1(M, \mathbb{R})$.

By definition, an element of $\bar A_{L,c}$ is an equivalence class of elements of $A_{L,c}$, for the equivalence relation $\sim$, where $(m_0, \tau_0) \sim (m_1, \tau_1)$ means that $\rho_{L,c}((m_0, \tau_0), (m_1, \tau_1)) = 0$. Here, $m_i \in M$, $\tau_i \in \mathbb{T}$, and $(m_i, \tau_i) \in A_{L,c}$, for $i = 0, 1$. Thus, for $X \in \bar A_{L,c}$, we have $X \subset A_{L,c} \subset M \times \mathbb{T}$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

If $\gamma : \mathbb{R} \to M$ is a continuous curve, we define the *α-limit set* $\alpha(\gamma)$ (respectively *ω-limit set* $\omega(\gamma)$) to be $\{(m, \tau) \in M \times \mathbb{T}$: for any neighborhood U of (m, τ) in $M \times \mathbb{T}$ and any $T > 0$, there exists $t < -T$ (respectively $t > T$) such that $(\gamma(t), t \mod 1) \in U\}$. Clearly, $\alpha(\gamma)$ and $\omega(\gamma)$ are closed sub-sets of $M \times \mathbb{T}$.

In [**Mat2**], we proved that if $\gamma : \mathbb{R} \to M$ is a c-minimizer, then there exist $X, Y \in \bar{A}_{L,c}$ such that $\alpha(\gamma) \subset X$ and $\omega(\gamma) \subset Y$. Moreover, γ is a regular c-minimizer if and only if $X = Y$. In addition, if $(m_0, \tau_0) \in X$, $(m_1, \tau_1) \in Y$ $(m_i \in M, \tau_i \in \mathbb{T}, i = 0, 1)$, and η is a Mañé critical closed 1-form such that $[\eta]_M = c$, then $h_\eta((m_0, \tau_0), (m_1, \tau_1)) = h_\eta((m_0, \tau_0), (\gamma(t), t \mod 1)) + h_\eta((\gamma(t), t \mod 1), (m_1, \tau_1))$. For brevity, we define $[\alpha(\gamma)], [\omega(\gamma)] \in \bar{A}_{L,c}$ to be the images of $\alpha(\gamma)$ and $\omega(\gamma)$ under the projection $A_{L,c} \to \bar{A}_{L,c}$, i.e. $[\alpha(\gamma)] = X$ and $[\omega(\gamma)] = Y$. We say *γ connects X to Y*.

In this section, we discuss the question, given $X, Y \in \bar{A}_{L,c}$, when does there exist a c-minimizer that connects X to Y?

In the case that $X = Y$, any c-minimizer γ in X (and there is always at least one) connects X to itself. Conversely, any c-minimizer that connects X to itself is regular and lies in X.

In the case when $X \neq Y$, there does not always exist a c-minimizer that connects X to Y, but nonetheless it is possible to connect X to Y in a generalized sense, as described in the Proposition below.

If S is a set and $\leq$ is a relation on S, we will say that $\leq$ is an *order* if:

- $x \leq y$ and $y \leq z$ implies $x \leq z$;
- $x \leq x$;
- $x \leq y$ and $y \leq x$ implies $x = y$; and
- for any x, y, we have $x \leq y$ or $y \leq x$.

We write $x < y$ to mean $x \leq y$ and $x \neq y$. We say y is *between* x and z if $x < y < z$. We say x is a *predecessor* of y and y is a *successor* of x relative to S if $x < y$ and there is no z between x and y.

PROPOSITION. *Let $X, Y \in \bar{A}_{L,c}$, $X \neq Y$. There exists a sub-set $\mathfrak{Z} = \mathfrak{Z}_{X,Y}$ of $\bar{A}_{L,c}$ and an order $\leq$ on $\mathfrak{Z}$ such that the following hold.*

- *X is the least and Y the greatest element of $\mathfrak{Z}$.*
- *The topology on $\mathfrak{Z}$ associated to the metric $\rho_{L,c}$ is the same as the topology associated to the order.*
- *$\mathfrak{Z}$ is compact with respect to this topology.*
- *Let $U, U', U^* \in \mathfrak{Z}$ with $U < U' < U^*$. Let $(m, \tau) \in U, (m', \tau') \in U'$, $(m^*, \tau^*) \in U^*$, where $m, m', m^* \in M$ and $\tau, \tau', \tau^* \in \mathbb{T}$. Then*

$$h_\eta((m, \tau), (m', \tau')) + h_\eta((m', \tau'), (m^*, \tau^*)) = h_\eta((m, \tau), (m^*, \tau^*)).$$

- *Let $U, V \in \mathfrak{Z}$. If V is a successor of U relative to $\mathfrak{Z}$, then there exists a c-minimizer $\gamma = \gamma_{U,V}$ such that $[\alpha(\gamma)] = U$ and $[\omega(\gamma)] = V$.*

Proof. Let η be a Mañé critical closed 1-form on $M \times \mathbb{T}$ such that $[\eta]_M = c$. Let $(m_0, \tau_0) \in X \cap M_{L,c}$, $(m_1, \tau_1) \in Y \cap M_{L,c}$. By definition (§1) of $h_\eta = h_{L,\eta}$, there exists a sequence of absolutely continuous curves $\gamma_i : [a_i, b_i] \to M$, $i = 1, 2, \ldots$, such that

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

$\gamma_i(a_i) = m_0, \gamma_i(b_i) = m_1, a_i \equiv \tau_0 \pmod 1, b_i \equiv \tau_1 \pmod 1, b_i - a_i \to +\infty$, and

$$\int_{a_i}^{b_i} (L - \widehat{\eta})(\gamma_i(t), \dot{\gamma}_i(t), t)\, dt \to h_\eta((m_0, \tau_0), (m_1, \tau_1)),$$

as $i \to +\infty$. By Tonelli's Theorem, there exists an absolutely continuous curve $\gamma_i^* : [a_i, b_i] \to M$ such that $\gamma_i^*(a_i) = \gamma_i(a_i)$, $\gamma_i^*(b_i) = \gamma_i(b_i)$ and γ_i^* minimizes $\int_{a_i}^{b_i} (L - \widehat{\eta})(\gamma_i^*(t), \dot{\gamma}_i^*(t), t)\, dt$ subject to the boundary condition. From the hypothesis that the Euler–Lagrange flow is complete and the minimizing property of γ_i^*, it follows that γ_i^* is C^2 and satisfies the Euler–Lagrange equation.

Let $\Gamma_i = \{(\gamma_i^*(t), \dot{\gamma}_i^*(t), t \mod 1) : a_i \le t \le b_i\} \subset TM \times \mathbb{T}$. Clearly Γ_i is compact. Moreover, it follows from the proof of [**Mat1**, Proposition 4] that $\bigcup_i \Gamma_i$ is relatively compact in $TM \times \mathbb{T}$. It follows that the sequence $\Gamma_1, \Gamma_2, \ldots$ has a subsequence that is convergent in the Hausdorff topology, and the limit Ω is a compact sub-set of $TM \times \mathbb{T}$. Without loss of generality, we may suppose that the sequence $\Gamma_1, \Gamma_2, \ldots$ converges to Ω in the Hausdorff topology.

Let $\gamma : \mathbb{R} \to M$ be an Euler–Lagrange trajectory (for L) such that, for some $t_0 \in \mathbb{R}$, we have $(\gamma(t_0), \dot{\gamma}(t_0), t_0 \bmod 1) \in \Omega$. Then either $(\gamma(t), \dot{\gamma}(t), t \bmod 1) \in \Omega$ for all $t \le t_0$ or there exists $a \le t_0$ such that $(\gamma(a), a \bmod 1) = (m_0, \tau_0)$ and $(\gamma(t), \dot{\gamma}(t), t \bmod 1) \in \Omega$ for all $a \le t \le t_0$. This holds because Ω is a limit of the Γ_i's in the Hausdorff topology and the same property holds for the Γ_i's. For the same reason, either $(\gamma(t), \dot{\gamma}(t), t \bmod 1) \in \Omega$ for all $t \ge t_0$ or there exists $b \ge t_0$ such that $(\gamma(b), b \bmod 1) = (m_1, \tau_1)$ and $(\gamma(t), \dot{\gamma}(t), t \bmod 1) \in \Omega$ for all $t_0 \le t \le b$.

Furthermore, γ is a c-minimizer. To prove this, we consider two cases.

The first case that we consider is when $(\gamma(t), t \bmod 1) \ne (m_i, \tau_i)$ for all $t \in \mathbb{R}$ and for $i = 0$ or 1. In this case, $(\gamma(t), \dot{\gamma}(t), t \bmod 1) \in \Omega$, for all $t \in \mathbb{R}$. To prove that γ is a c-minimizer, we suppose the contrary and obtain a contradiction.

Thus, we suppose that there exists $a < b, c < d$ such that $c - a \in \mathbb{Z}, d - b \in \mathbb{Z}$, and an absolutely continuous curve $\widetilde{\gamma} : [c, d] \to M$ such that $\widetilde{\gamma}(c) = \gamma(a), \widetilde{\gamma}(d) = \gamma(b)$, and

$$\int_c^d (L - \widehat{\eta})(\widetilde{\gamma}, \dot{\widetilde{\gamma}}, t) < \int_a^b (L - \widehat{\eta})(\gamma, \dot{\gamma}, t)\, dt.$$

In view of the facts that $\Omega = \lim_{i \to +\infty} \Gamma_i$ and $\{(\gamma(t), \dot{\gamma}(t), t \bmod 1) : t \in \mathbb{R}\} \subset \Omega$, it follows that for large i there exist $a_i^* \in [a_i, b_i]$ with $a_i^* - a \in \mathbb{Z}$ and $\gamma_i^*(a_i^*) \to \gamma(a), \dot{\gamma}_i^*(a_i^*) \to \dot{\gamma}(a)$, as $i \to +\infty$. Here, we use the fact that $(\gamma(a), \dot{\gamma}(a), a \bmod 1) \in \Omega$.

It follows that $\gamma_i^*(t + a_i^* - a) \to \gamma(t)$ and $\dot{\gamma}_i^*(t + a_i^* - a) \to \dot{\gamma}(t)$ uniformly on any compact interval, as $i \to +\infty$. Hence, there exists $\widetilde{\gamma}_i^* : [a_i, b_i + d - c - b + a] \to M$ such that $\widetilde{\gamma}_i^*(t) = \gamma_i^*(t)$ for $a_i \le t \le a_i^*$, $\widetilde{\gamma}_i^*(t + d - c - b + a) = \gamma_i^*(t)$ for $b_i \ge t \ge a_i^* + b - a$, and $\widetilde{\gamma}_i^*(t + a_i^* - c) \to \widetilde{\gamma}(t)$ and $\dot{\widetilde{\gamma}}_i^*(t + a_i^* - c) \to \dot{\widetilde{\gamma}}(t)$ uniformly on $[c, d]$, as $i \to +\infty$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Then

$$\lim_{i\to+\infty}\int_{a_i}^{b_i+d-c-a+b}(L-\widehat{\eta})(\widetilde{\gamma}_i^*,\dot{\widetilde{\gamma}}_i^*,t)\,dt$$
$$=\lim_{i\to+\infty}\int_{a_i}^{b_i}(L-\widehat{\eta})(\gamma_i^*,\dot\gamma_i^*,t)\,dt+\int_c^d(L-\widehat{\eta})(\widetilde{\gamma},\dot{\widetilde{\gamma}},t)\,dt$$
$$-\int_a^b(L-\widehat{\eta})(\gamma_1\dot\gamma,t)\,dt<h_\eta((m_0,\tau_0),(m_1,\tau_1)).$$

This contradicts the definition of $h_\eta((m_0,\tau_0),(m_1,\tau_1))$, and thereby proves that γ is a c-minimizer, in the first case.

In the other case, $(\gamma(\hat t_0),\ \hat t_0\mod 1)=(m_i,\tau_i)$ for some $\hat t_0\in\mathbb{R}$ and for $i=0$ or 1. We will give the proof that γ is a c-minimizer in the sub-case when $i=0$. The other sub-case is similar.

Since $(m_0,\tau_0)\in M_{L,c}$, there exists a unique $v_0\in TM_{m_0}$ such that $(m_0,v_0,\tau_0)\in\dot M_{L,c}$, by the graph theorem. If $\dot\gamma(\hat t_0)=v_0$, then γ is a c-minimizer, so we may suppose that $\dot\gamma(\hat t_0)\neq v_0$. Under this assumption, we will obtain a contradiction by a curve shortening argument, as follows.

For an arbitrarily small neighborhood U of m_0 in M, there exists a c-minimizer $\mu:\mathbb{R}\to M_{L,c}$ and $t_{-1}<\hat t_0$, $t_{-1}\equiv\hat t_0 \mod 1$ with $\mu(t_{-1})\in U$ and $\mu(t_0)\in U$. This is a consequence of Poincaré recurrence applied to the Euler–Lagrange flow restricted to $\dot M_{L,c}$. Then

$$\int_{t_{-1}}^{\hat t_0}(L-\widehat{\eta})(\mu,\dot\mu,t)\,dt=h_{L,c}((\mu(t_{-1}),t_{-1}),(\mu(\hat t_0),\hat t_0)),$$

since $\mu(t)\in M_{L,c}$ for all $t\in\mathbb{R}$. Note that, as U shrinks to m_0, the right-hand side tends to zero, since $(m_0,\tau_0)\in M_{L,c}$ and $\mu(t_{-1}),\mu(\hat t_0)\in U$.

Let $\widehat\mu:\mathbb{R}\to M$ be the unique Euler–Lagrange trajectory such that $\widehat\mu(\hat t_0)=m_0$ and $\dot{\widehat\mu}(\hat t_0)=v_0$. Let δ be a small positive number and let $\widehat\gamma$ (respectively $\widehat\gamma^*$): $[\hat t_0-\delta,\hat t_0+\delta]\to M$ minimize $\int_{\hat t_0-\delta}^{\hat t_0+\delta}(L-\widehat{\eta})(\widehat\gamma,\dot{\widehat\gamma},t)\,dt$ subject to the boundary conditions $\widehat\gamma(t_0-\delta)=\widehat\mu(\hat t_0-\delta)$ (respectively $\widehat\gamma^*(\hat t_0-\delta)=\mu(\hat t_0-\delta)$) and $\widehat\gamma(\hat t_0+\delta)=\gamma(\hat t_0+\delta)$ (respectively $\widehat\gamma^*(\hat t_0+\delta)=\gamma(\hat t_0+\delta)$). Because $\dot\gamma(\hat t_0)\neq v_0=\dot{\widehat\mu}(\hat t_0)$, we have

$$\int_{\hat t_0-\delta}^{\hat t_0+\delta}(L-\widehat{\eta})(\widehat\gamma,\dot{\widehat\gamma},t)\,dt+\Delta<\int_{\hat t_0-\delta}^{\hat t_0}(L-\widehat{\eta})(\widehat\mu,\dot{\widehat\mu},t)\,dt+\int_{\hat t_0}^{\hat t_0+\delta}(L-\widehat{\eta})(\gamma,\dot\gamma,t)\,dt,\quad(\star)$$

for some positive Δ.

Let $\gamma^*:[t_{-1},\infty)\to M$ be such that:

- $\gamma^*|[t_{-1},t_{-1}+\delta]$ minimizes $\int_{t_{-1}}^{t_{-1}+\delta}(L-\widehat{\eta})(\gamma^*,\dot\gamma^*,t)\,dt$ subject to the boundary conditions $\gamma^*(t_{-1})=m_0$, $\gamma^*(t_{-1}+\delta)=\mu(t_{-1}+\delta)$;
- $\gamma^*(t)=\mu(t)$, for $t_{-1}+\delta\le t\le\hat t_0-\delta$;
- $\gamma^*(t)=\widehat\gamma^*(t)$, for $t_0-\delta\le t\le t_0+\delta$; and
- $\gamma^*(t)=\gamma(t)$, for $t_0+\delta\le t$.

Just as in the previous case, for large i there exist $a_i^*\in[a_i,b_i]$ with $a_i^*-\hat t_0\in\mathbb{Z}$ and $\gamma_i^*(a_i^*)\to\gamma_i(\hat t_0)=m_0$, $\dot\gamma_i^*(a_i^*)\to\dot\gamma(\hat t_0)$, as $i\to+\infty$. It follows that $\gamma_i^*(t+a_i^*-\hat t_0)\to\gamma(t)$ and $\dot\gamma_i^*(t+a_i^*-\hat t_0)\to\dot\gamma(t)$ uniformly on any compact sub-interval of $[\hat t_0,+\infty)$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Let $\widetilde{\gamma}_i^*\colon [t_{-1}, b_i] \to M$ be such that $\widetilde{\gamma}_i^*(t) = \gamma_i^*(t)$ for $a_i^* + \delta \le t \le b_i$ and $\widetilde{\gamma}_i^*(t + a_i^* - \hat{t}_0) \to \gamma^*(t)$, $\dot{\widetilde{\gamma}}_i^*(t + a_i^* - \hat{t}_0) \to \dot{\gamma}^*(t)$ uniformly on any compact sub-interval of $[t_{-1}, +\infty)$. If U is small enough, we have

$$\int_{t_{-1}}^{b_i} (L - \widehat{\eta})(\widetilde{\gamma}_i^*, \dot{\widetilde{\gamma}}_i^*, t)\, dt \le \int_{a_i}^{b_i} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}_i^*, t)\, dt - \Delta. \tag{$\dagger$}$$

for all sufficiently large i. This follows from ($\star$) and the fact that if U is very small, then $\mu(t_{-1})$ and $\mu(\hat{t}_0)$ are very near to m_0, and hence

$$\left| \int_{t_{-1}}^{\hat{t}_0} (L - \widehat{\eta})(\mu, \dot{\mu}, t)\, dt \right|,$$

$$\left| \int_{t_{-1}}^{t_{-1}+\delta} (L - \widehat{\eta})(\gamma^*, \dot{\gamma}^*, t)\, dt - \int_{t_{-1}}^{t_{-1}+\delta} (L - \widehat{\eta})(\mu, \dot{\mu}, t)\, dt \right|,$$

$$\left| \int_{t_0-\delta}^{\hat{t}_0} (L - \widehat{\eta})(\hat{\mu}, \dot{\hat{\mu}}, t)\, dt - \int_{\hat{t}_0-\delta}^{\hat{t}_0} (L - \widehat{\eta})(\mu, \dot{\mu}, t)\, dt \right|,$$

and

$$\left| \int_{\hat{t}_0-\delta}^{\hat{t}_0+\delta} (L - \widehat{\eta})(\widehat{\gamma}^*, \dot{\widehat{\gamma}}, t) - \int_{\hat{t}_0-\delta}^{\hat{t}_0+\delta} (L - \widehat{\eta})(\gamma^*, \gamma^*, t)\, dt \right|$$

are very small. Indeed, these quantities can be made arbitrarily small by taking U small enough. Note that the choices are made in the following order: first, we choose $\delta > 0$ (this may be chosen arbitrarily), then U small enough so that ($\dagger$) follows from ($\star$), then $t_{-1} \ll \hat{t}_0$ so that there is a Euler–Lagrange trajectory with $\mu(\hat{t}_0), \mu(t_1) \in U$. Note also that, since $\mu(\hat{t}_0)$ is near m_0, so is $\dot{\mu}(\hat{t}_0)$ near v_0, so $\mu(t), \dot{\mu}(t)$ are near $\widehat{\mu}(t)$, $\dot{\widehat{\mu}}(t)$ for $\hat{t}_0 - \delta \le t \le \hat{t}_0$.

From ($\dagger$), it follows that

$$\begin{aligned}\liminf_{i\to+\infty} \int_{t_{-2}}^{b_i} (L - \widehat{\eta})(\widetilde{\gamma}_i^*, \dot{\widetilde{\gamma}}_i^*, t)\, dt &\le \lim_{i\to\infty} \int_{a_i}^{b_i} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}^*, t)\, dt - \Delta \\ &= h_{L,c}((m_0, \hat{t}_0), (m_1, t_1)) - \Delta.\end{aligned}$$

In fact, we can choose (if necessary) $t_{-1} = t_{-1}^i$ to depend on i, so that $|b_i - t_1^i| \to +\infty$ as $i \to +\infty$. Then we have a contradiction to the definition of $h_{L,c}((m_0, \tau_0), (m_1, \tau_1))$. This contradiction shows that we cannot have $\dot{\gamma}(\hat{t}_0) \ne v_0$. Hence γ is c-minimal in this case, too.

Thus, we have proved that if $(\gamma(\hat{t}_0), \hat{t}_0 \mod 1) = (m_0, \tau_0)$, then γ is a c-minimizer. The proof that if $(\gamma(\hat{t}_0), \hat{t}_0 \mod 1) = (m_1, \tau_1)$, then γ is a c-minimizer is similar.

We define a relation $\le$ on Ω as follows.

Let $(m, v, \tau), (m', v', \tau') \in \Omega$ (where $m, m' \in M, v \in TM_m, v' \in TM_{m'}$, and $\tau, \tau' \in \mathbb{T}$). We say that $(m, v, \tau) \le (m', v', \tau')$ if for every neighborhood N of (m, v, τ) and every neighborhood N' of (m', v', τ') in $TM \times \mathbb{T}$, there exists i and $t_0 \le t_1$ in $[a_i, b_i]$ such that $(\gamma_i^*(t_0), \dot{\gamma}_i^*(t_0), t_0 \mod 1) \in N$ and $(\gamma_i^*(t_1), \dot{\gamma}_i^*(t_1), t_1 \mod 1) \in N'$.

We let $\mathfrak{Z}$ denote the image of $\Omega \cap \dot{A}_{L,c}$ under the projection $\dot{A}_{L,c} \to \bar{A}_{L,c}$ (i.e. the composition of the projections $\dot{A}_{L,c} \to A_{L,c}$ and $A_{L,c} \to \bar{A}_{L,c}$).

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

We define a relation $\leq$ on $\mathfrak{Z}$ as follows.

Let $U, U' \in \mathfrak{Z}$. We say that $U \leq U'$ if there exist $(m, \tau) \in U, (m', \tau') \in U'$, $v \in TM_m$, $v' \in TM_{m'}$ such that $(m, v, \tau) \leq (m', v', \tau')$.

The relation $\leq$ on Ω is not necessarily an order. On the other hand, we prove below that the relation $\leq$ on $\mathfrak{Z}$ is an order.

By definition $\Omega = \lim_{i \to +\infty} \Gamma_i$ (convergence in the Hausdorff topology). It follows that, for any $(m, v, \tau), (m', v', \tau') \in \Omega$, at least one of the relations $(m, v, \tau) \leq (m', v', \tau')$ or $(m', v'\tau') \leq (m, v, \tau)$ holds. Consequently, for any $U, U' \in \mathfrak{Z}$, at least one of the relations $U \leq U'$ or $U' \leq U$ holds.

LEMMA. *Let (m, v, τ), $(m', v', \tau') \in \Omega$ and suppose that at least one of these is an element of $\dot{A}_{L,c}$. Suppose that, for each positive integer i, there exist real numbers t_i, t_i' such that $a_i \leq t_i \leq t_i' \leq b_i$ and*

$$(\gamma_i^*(t_i), \dot{\gamma}_i^*(t_i), \ t_i \mod 1) \to (m, v, \tau),$$
$$(\gamma_i^*(t_i'), \dot{\gamma}_i^*(t_i'), \ t_i' \mod 1) \to (m', v', \tau'),$$

as $i \to +\infty$. Then

$$\int_{t_i}^{t_i'} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}_i^*, t)\, dt \to h_\eta((m, \tau), (m', \tau')),$$

as $i \to +\infty$.

Proof. Suppose that there exists $\Delta > 0$, such that

$$\limsup_{i \to +\infty} \int_{t_i}^{t_i'} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}_i^*, t)\, dt > h_\eta((m, \tau)(m', \tau')) + \Delta.$$

From the definition of $h_\eta((m, \tau), (m', \tau'))$, it follows that for infinitely many i there exist $\tilde{t}_i, \tilde{t}_i' \in \mathbb{R}$, with $\tilde{t}_i - t_i, \tilde{t}_i' - t_i' \in \mathbb{Z}$, and $\tilde{\gamma}_i : [\tilde{t}_i, \tilde{t}_i'] \to M$ such that

$$\int_{\tilde{t}_i}^{\tilde{t}_i'} (L - \widehat{\eta})(\tilde{\gamma}_i, \dot{\tilde{\gamma}}_i, t)\, dt \leq \int_{t_i}^{t_i'} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}_i^*, t)\, dt - \Delta,$$

and $\tilde{\gamma}_i(\tilde{t}_i) = \gamma_i^*(t_i)$, $\tilde{\gamma}_i(\tilde{t}_i') = \gamma_i^*(t_i')$. Moreover, we may choose $\tilde{t}_i' - \tilde{t}_i$ to be arbitrarily large. We extend $\tilde{\gamma}_i$ to $[a_i + \tilde{t}_i - t_i, b_i + \tilde{t}_i' - t_i']$ by setting $\tilde{\gamma}_i(t + \tilde{t}_i - t_i) = \gamma_i^*(t)$ for $a_i \leq t \leq t_i$ and $\tilde{\gamma}_i(t + \tilde{t}_i' - t_i') = \gamma_i^*(t)$ for $t_i' \leq t \leq b_i$. For such i, we have

$$\int_{a_i + \tilde{t}_i - t_i}^{b_i + \tilde{t}_i' - t_i'} (L - \widehat{\eta})(\tilde{\gamma}_i, \dot{\tilde{\gamma}}_i, t)\, dt \leq \int_{a_i}^{b_i} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}_i^*, t)\, dt - \Delta.$$

Since

$$\lim_{i \to +\infty} \int_{a_i}^{b_i} (L - \widehat{\eta})(\gamma_i^*, \dot{\gamma}_i^*, t)\, dt = h_\eta((m_0, \tau_0), (m_1, \tau_1)),$$

it follows that

$$\liminf_{i \to +\infty} \int_{a_i + \tilde{t}_i - t_i}^{b_i + \hat{t}_i' - t_i'} (L - \widehat{\eta})(\tilde{\gamma}_i, \dot{\tilde{\gamma}}_i, t)\, dt \leq h_\eta((m_0, \tau_0), (m_1, \tau_1)) - \Delta,$$

which contradicts the definition of $h_\eta((m_0, \tau_0), (m_1, \tau_1))$.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

Hence,

$$\limsup_{i\to+\infty}\int_{t_i}^{t_i'}(L-\widehat{\eta})(\gamma_i^*,\dot\gamma_i^*,t)\,dt\le h_\eta((m,\tau),(m',\tau')).$$

Now, suppose that there exists $\Delta>0$ such that

$$\liminf_{i\to+\infty}\int_{t_i}^{t_i'}(L-\widehat{\eta})(\gamma_i^*,\dot\gamma_i^*,t)\,dt< h_\eta((m,\tau),(m',\tau'))-\Delta.$$

In this case, there exist $\tilde a<\tilde b$ and $\tilde\gamma:[\tilde a,\tilde b]\to M$ such that $\tilde a\equiv\tau \mod 1$, $\tilde b\equiv\tau' \mod 1$, $\widetilde\gamma(\tilde a)=m$, $\widetilde\gamma(\tilde b)m'$, and

$$\int_{\tilde a}^{\tilde b}(L-\widehat{\eta})(\widetilde\gamma,\dot{\widetilde\gamma},t)\,dt< h_\eta((m,\tau),(m',\tau')-\Delta. \qquad (\star)$$

Suppose that $(m,v,\tau)\in\dot A_{L,c}$. (The case that $(m',v',\tau')\in\dot A_{L,c}$ is similar.) Then $(m,\tau)\in A_{L,c}$, so $h_\eta((m,\tau),(m,\tau))=\rho_c((m,\tau),(m,\tau))/2=0$. For any δ, $T_0>0$, it is possible to find an integer $T\ge T_0$ and an absolutely continuous curve $\widehat\gamma:[\tilde a-T,\tilde a]\to M$ such that $\widehat\gamma(\tilde a-T)=\widehat\gamma(\tilde a)=m$ and

$$0\le\int_{\tilde a-T}^{\tilde a}(L-\widehat{\eta})(\widehat\gamma,\dot{\widehat\gamma},t)\,dt<\delta.$$

Extend $\widetilde\gamma$ to $[\tilde a-T,\tilde b]$ by setting $\widetilde\gamma(t)=\widehat\gamma(t)$ for $\tilde a-T\le t\le\tilde a$. If δ is small enough, the inequality $(\star)$ is still valid. However, since T_0 can be chosen as large as we like, this contradicts the definition of $h_\eta((m,\tau),(m',\tau'))$. □

Now, we continue with the proof of the Proposition.

Suppose that $(m,v,\tau),(m',v',\tau')\in\Omega$ and $(m,v,\tau)\le(m',v',\tau')$. The Lemma implies that

$$h_\eta((m_0,\tau_0),(m,\tau))+h_\eta((m,\tau),(m',\tau'))=h_\eta((m_0,\tau_0),(m',\tau')).$$

(To apply the Lemma, it may be necessary to replace γ_i^* with a subsequence.)

If, in addition, we have $(m',v',\tau')\le(m,v,\tau)$, then

$$h_\eta((m_0,\tau_0),(m',\tau'))+h_\eta((m',\tau'),(m,\tau))=h_\eta((m_0,\tau_0),(m,\tau)).$$

Adding these two equations and cancelling, we obtain

$$\rho_c((m,\tau),(m',\tau'))=0.$$

Thus, we have shown that if both $(m,v,\tau)\le(m',v',\tau')$ and $(m',v',\tau')\le(m,v,\tau)$, then the ρ_c-pseudo-distance between (m,τ) and (m',τ') vanishes. Consequently, if $U,U'\in\mathfrak{Z}$ and $U\le U'$ and $U'\le U$, then $U=U'$.

Suppose that $(m,v,\tau),(m',v',\tau')\in\Omega$ and the relation $(m',v',\tau')\le(m,v,\tau)$ does not hold. Let N and N' be neighborhoods of (m,v,τ) and (m',v',τ'), respectively, in $TM\times\mathbb{T}$. Since $\Omega=\lim_{i\to+\infty}\Gamma_i$, there exists t_0^i,t_1^i in $[a_i,b_i]$ such that $(\gamma_i^*(t_0^i),\dot\gamma_i^*(t_0^i),t_0^i \mod 1)\in N$ and $(\gamma_i^*(t_1^i),\dot\gamma^*),(t_1^i),t_1^i \mod 1)\in N'$, for all large i. Since $(m',v',\tau')\le(m,v,\tau)$ does not hold, we have $t_0^i<t_1^i$ if N and N' are small enough (and i is large enough).

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

This holds, in particular, if $(m, v, \tau) \leq (m', v', \tau')$ and $\rho_c((m, \tau), (m', \tau')) > 0$. It follows that if $(m, v, \tau) \leq (m', v', \tau')$ and $(m', v', \tau') \leq (m^*, v^*, \tau^*)$, then $(m, v, \tau) \leq (m^*, v^*, \tau^*)$, provided that $\rho_c((m, \tau), (m', \tau')) > 0$. Consequently, the relation $\leq$ on $\mathfrak{Z}$ is transitive.

We have thus shown that the relation $\leq$ on $\mathfrak{Z}$ is an order. (However, the relation $\leq$ on Ω is not necessarily an order.)

It follows immediately from the definitions that X is the least and Y the greatest element of $\mathfrak{Z}$.

Since Ω is obviously compact and $\mathfrak{Z}$ is a quotient space of Ω, it follows that $\mathfrak{Z}$ is compact with respect to the metric topology (the topology associated to $\rho_{L,c}$).

Let $U \in \mathfrak{Z}$. Next, we show that $\{U^* \in \mathfrak{Z} : U^* \leq U\}$ is closed with respect to the metric topology on $\mathfrak{Z}$. For, let $U_i \leq U$, for $i = 1, 2, \ldots$. Suppose that $U_i \to U^*$, with respect to the metric topology, as $i \to +\infty$. Suppose that $U^* > U$. Let $(m_i, \tau_i) \in U_i$, $(m, \tau) \in U$, $(m^*, \tau^*) \in U^*$, $m_i, m, m^* \in M$, $\tau_i, \tau, \tau^* \in \mathbb{T}$. By passing to a subsequence and choosing $(m_i, \tau_i) \in U_i$ appropriately, we may suppose that $(m_i, \tau_i) \to (m^*, \tau^*)$, as $i \to +\infty$. Then

$$\begin{aligned} 0 &= \lim_{i \to +\infty} h_\eta((m_i, \tau_i), (m^*, \tau^*)) \\ &= \lim_{i \to +\infty} h_\eta((m_i, \tau_i), (m, \tau)) + h_\eta((m, \tau), (m^*, \tau^*)) \\ &= \rho_c((m, \tau), (m^*, \tau^*)) > 0, \end{aligned}$$

a contradiction. Here, the second equation is a consequence of the assumption $U_i \leq U < U^*$ and the Lemma.

This contradiction shows that $U^* \leq U$. Hence, $\{U^* \in \mathfrak{Z} : U^* \leq U\}$ is closed with respect to the metric topology. Similarly, $\{U^* \in \mathfrak{Z} : U \leq U^*\}$ is closed with respect to the metric topology. If follows that any closed interval in $\mathfrak{Z}$ is closed with respect to the metric topology. Taking complements, we obtain that any open interval in $\mathfrak{Z}$ is open with respect to the metric topology. It follows that the identity mapping ($\mathfrak{Z}$, metric topology) $\to$ ($\mathfrak{Z}$, order topology) is continuous. Since ($\mathfrak{Z}$, metric topology) is compact, it then follows from an elementary result in general topology that id: ($\mathfrak{Z}$, metric topology) $\to$ ($\mathfrak{Z}$, order topology) is a homeomorphism, i.e. the two topologies are the same.

Let $U, U', U^* \in \mathfrak{Z}$ with $U < U' < U^*$. Let $(m, \tau) \in U$, $(m', \tau') \in U'$, $(m^*, \tau^*) \in U^*$. The equation

$$h_\eta((m, \tau), (m', \tau')) + h_\eta((m', \tau'), (m^*, \tau^*)) = h_\eta((m, \tau), (m^*, \tau^*))$$

follows immediately from the Lemma.

Let $U, V \in \mathfrak{Z}$ and suppose V is a successor of U. Let $\Omega_0 = \{(m, v, \tau) \in \Omega$: there exists $(m', v', \tau') \in \Omega$, $(m', \tau') \in U$ such that $(m, v, \tau) \leq (m', v', \tau')\}$. Let $\Omega_1 = \{(m, v, \tau) \in \Omega$: there exists $(m', v', \tau') \in \Omega$, $(m', v') \in V$ such that $(m', v', \tau') \leq (m, v, \tau)\}$. The argument above that showed that $\{U^* \in \mathfrak{Z} : U^* \leq U\}$ is closed in the metric topology also shows that Ω_0 is closed. Similarly Ω_1 is closed. Obviously, Ω_0 and Ω_1 are mutually disjoint.

It follows easily from the definition of Ω that it is connected. Since Ω_0 and Ω_1 are closed and mutually disjoint, it follows that there exists $(m, v.\tau) \in \Omega \setminus (\Omega_0 \cup \Omega_1)$. Let $t_0 \equiv \tau$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

(mod 1) and let $\gamma : \mathbb{R} \to M$ be the Euler–Lagrange trajectory such that $\gamma(t_0) = m$, $\dot\gamma(t_0) = v$. We proved previously that γ is c-minimal.

Moreover, either $(\gamma(t), \dot\gamma(t), t \mod 1) \in \Omega$ for all $t \le t_0$ or there exists $a \le t_0$ such that $(\gamma(a), a \mod 1) = (m_0, \tau_0)$. Since γ is c-minimal, $(\gamma(t), t \mod 1) \in X$ for all t in the latter case. This implies that $(\gamma(t), \dot\gamma(t), t \mod 1) \in \Omega_0$, contrary to $(\gamma(t_0), \dot\gamma(t_0), t_0 \mod 1) = (m, v, \tau) \notin \Omega_0$. This contradiction shows that $(\gamma(t), \dot\gamma(t), t \mod 1) \in \Omega$ for all $t \le t_0$. A similar argument shows that $(\gamma(t), \dot\gamma(t), t \mod 1) \in \Omega$ for all $t \ge t_0$.

It is easy to see that there exists $(m', \tau') \in \alpha(\gamma)$ and $v' \in TM_{m'}$, $(m', \tau', v') \in \Omega$, such that $(m', v', \tau') < (m, v, \tau)$. Hence, $[\alpha(\gamma)] < V$. Since U is a predecessor of V, this implies that $[\alpha(\gamma)] \le U$. Similarly, $[\omega(\gamma)] \ge V$.

Suppose $[\alpha(\gamma)] < U$. There exist $a_i \le t_i, t_i', t_i^* \le b_i$ such that $(\gamma_i^*(t_i), \dot\gamma_i^*(t_i), t_i \mod 1) \to (m, v, \tau)$, $(\gamma_i^*(t_i'), \dot\gamma_i^*(t_i'), t_i' \mod 1) \to (m', v', \tau') \in U$, and $(\gamma_i^*(t_i^*), \dot\gamma_i^*(t_i^*), t_i^* \mod 1) \to (m^*, v^*, \tau^*) \in [\alpha(\gamma)]$. The existence of such sequences is an immediate consequence of the definition of Ω and the facts $(m, v, \tau) \in \Omega$, $U \subset \Omega$, $[\alpha(\gamma)] \subset U$. If $t_i' \le t_i^*$ for infinitely many i, we have $U \le [\alpha(\gamma)]$. Since this is not so, it follows that $t_i^* < t_i'$ for all but finitely many i. Similarly, $t_i' < t_i$ for all but finitely many i. It then follows from the Lemma that

$$h_\eta((m^*, \tau^*), (m', \tau')) + h_\eta((m', \tau'), (m, \tau)) = h_\eta((m^*, \tau^*), (m, \tau)).$$

The only thing that we have used about (m, τ) in deriving this equation is that it has the form $(m, \tau) = (\gamma(t_0), t_0 \mod 1)$ for some $t_0 \in \mathbb{R}$. Thus, we may let t_0 tend to $-\infty$. Let $(\widetilde m, \tilde\tau)$ be a limit point of $(\gamma(t_0), t_0 \mod 1)$ as t_0 tends to $-\infty$. Taking limits, we obtain

$$h_\eta((m^*, \tau^*), (m', \tau')) + h_\eta((m', \tau'), (\widetilde m, \tilde\tau)) = h_\eta((m^*, \tau^*), (\widetilde m, \tilde\tau)).$$

Since $(\widetilde m, \tilde\tau) \in [\alpha(\gamma)]$, we have $h_\eta((m^*, \tau^*), (\widetilde m, \tilde\tau)) + h_\eta((\widetilde m, \tilde\tau), (m^*, \tau^*)) = \rho_c((m^*, \tau^*), (\widetilde m, \tilde\tau)) = 0$. Adding $h_\eta((\widetilde m, \tilde\tau), (m^*, \tau^*))$ to both sides of the equation above, we get

$$\begin{aligned} 0 &< \rho_c((m^*, \tau^*), (m', \tau')) \\ &\le h_\eta((m^*, \tau^*), (m', \tau')) + h_\eta((m', \tau'), (\widetilde m, \tilde\tau)) + h_\eta((\widetilde m, \tilde\tau), (m^*, \tau^*)) = 0. \end{aligned}$$

This contradiction shows that $[\alpha(\gamma)] = U$. Similarly, $[\omega(\gamma)] = V$.

17. *An example where $\dot M_{L,c} = \dot A_{L,c} = \dot N_{L,c}$ and $\bar A_{L,c}$ is not reduced to one point*

It is easy to produce examples where $\dot M_{L,c} = \dot A_{L,c} = \dot N_{L,c}$ and $\bar A_{L,c}$ is one point. For example, if $M = \mathbb{T}^1$, $\omega \in \mathbb{R} = H^1(\mathbb{T}^1; \mathbb{R})$ is irrational, and $c = \mathcal{L}_{\beta_L}(\omega)$, then for a generic choice of L, we have $\dot M_{L,c} = \dot A_{L,c} = \dot N_{L,c}$. Moreover, $\bar A_{L,c}$ is one point. See [**B**] or [**FoM**].

It is not easy to produce examples where $\dot A_{L,c} = \dot N_{L,c}$ and $\bar A_{L,c}$ is more than one point. In this section, we first discuss the difficulty and then give an example.

Definition 1. Let $\mathfrak{Z}$ be a compact sub-set of $\bar A_{L,c}$. By a *direction* on $\mathfrak{Z}$, we mean an order $\le$ with the following properties.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

- The order topology on $\mathfrak{Z}$ is the same as the metric topology (associated to $\rho_{L,c}$).
- Let $U, U', U^* \in \mathfrak{Z}$ with $U < U' < U^*$. Let $(m, \tau) \in U$, $(m', \tau') \in U'$, $(m^*, \tau^*) \in U^*$, where $m, m', m^* \in M$ and $\tau, \tau' \tau^* \in \mathbb{T}$. Then

$$h_\eta((m, \tau), (m', \tau')) + h_\eta((m', \tau'), (m^*, \tau^*)) = h_\eta((m, \tau), (m^*, \tau^*)).$$

(Throughout this section M is a compact manifold, L is a Lagrangian on $TM \times \mathbb{T}$ satisfying the usual conditions, $c \in H^1(M; \mathbb{R})$, and η is a Mañé critical 1-form on $M \times \mathbb{T}$ such that $[\eta]_M = c$.)

By a *directed sub-set* of $\bar{A}_{L,c}$, we mean a pair $(\mathfrak{Z}, \leq)$, where $\mathfrak{Z}$ is a compact sub-set of $\bar{A}_{L,c}$ and $\leq$ is a direction on $\mathfrak{Z}$.

Let $X, Y \in \bar{A}_{L,c}$, $X \neq Y$. The Proposition in §16 asserts that there is a directed sub-set $\mathfrak{Z} = \mathfrak{Z}_{X,Y}$ of $\bar{A}_{L,c}$ such that X is the least and Y the greatest element of $\mathfrak{Z}$ and such that if $U, V \in \mathfrak{Z}$ and V is a successor of U relative to $\mathfrak{Z}$, then there is a c-minimizer $\gamma = \gamma_{U,V}$ such that $[\alpha(\gamma)] = U$ and $[\omega(\gamma)] = V$. Since γ is a c-minimizer, we have that $\{(\gamma(t), \dot{\gamma}(t), t \mod 1) : t \in \mathbb{R}\}$ is a sub-set of $\dot{N}_{L,c}$. On the other hand, since $[\alpha(\gamma)] \neq [\omega(\gamma)]$, we have that $\{(\gamma(t), \dot{\gamma}(t), t \mod 1) : t \in \mathbb{R}\}$ is not a sub-set of $\dot{A}_{L,c}$. This shows that if $\mathfrak{Z}_{X,Y}$ contains U, V such that V is a successor of U relative to $\mathfrak{Z}$, then $\dot{A}_{L,c}$ is properly contained in $\dot{N}_{L,c}$.

To put this another way, the hypothesis that $\dot{A}_{L,c} = \dot{N}_{L,c}$ implies that $\mathfrak{Z}_{X,Y}$ is order isomorphic to a closed interval (with endpoints X and Y) for every $X, Y \in \bar{A}_{L,c}$. Finding L and c so that $\bar{A}_{L,c}$ has this property is the difficulty alluded to above.

The main result in this paper (Theorem 1 in §2) provides an example with such a property. Let L_0 be the Lagrangian whose existence is asserted in Theorem 1. Let $X, Y \in \bar{A}_{L_0,0}$. Since $\bar{A}_{L_0,0}$ is isometric to an interval, we may speak of the sub-interval $[X, Y]$ having X and Y as endpoints. The fact that $\bar{A}_{L_0,0}$ is isometric to an interval (for the metric $\rho_{L_0,0}$) implies that $[X, Y]$ satisfies the conditions given in the Proposition in §16 for both $\mathfrak{Z}_{X,Y}$ and $\mathfrak{Z}_{Y,X}$. (Note that $\mathfrak{Z}_{X,Y}$ is not necessarily uniquely determined. Uniqueness may fail even in the case that $\bar{A}_{L_0,0}$ is isometric to an interval.) In other words, there are two directions on $[X, Y]$, one the opposite of the other. Each direction is induced by an isometry of $[X, Y]$ with an interval.

To show this, we consider an isometry $\Phi : [X, Y] \to [a, b]$, where $a < b \in \mathbb{R}$. We suppose that $\Phi(X) = a$ and $\Phi(Y) = b$. We provide $[X, Y]$ with the order induced by Φ, i.e. for $U, V \in [X, Y]$, we have $U \leq V$ if and only if $\Phi(U) \leq \Phi(V)$.

Let $U, U', U^* \in \bar{A}_{L_0,0}$ and let $(m, \tau) \in U$, $(m', \tau') \in U'$, $(m^*, \tau^*) \in U^*$. Suppose $U < U' < U^*$. We have

$$h_\eta((m, \tau), (m^*, \tau^*)) \leq h_\eta((m, \tau), (m', \tau')) + h_\eta((m', \tau'), (m^*, \tau^*)) \qquad (\star)$$

and the inequality obtained by interchanging (m, τ) and (m^*, τ^*). These inequalities are an immediate consequence of the definition of h_η. We have

$$\rho_c((m, \tau), (m^*, \tau^*)) = \rho_c((m, \tau), (m', \tau')) + \rho_c((m', \tau'), (m^*, \tau^*)),$$

since $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval. In view of the definition of ρ_c, it follows from $(\star)$ and from the inequality obtained by interchanging (m, τ) and (m^*, τ^*) in $(\star)$ that $(\star)$ is an equality, not just an inequality.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

This shows that the order provided above for $[X, Y]$ is a direction. Similarly, the opposite order is also a direction.

The discussion so far shows that L_0 satisfies the necessary condition given above if $\dot{A}_{L_0,0} = \dot{N}_{L_0,0}$. In fact, some examples given by the Proposition in §5 satisfy $\dot{A}_{L_0,0} = \dot{N}_{L_0,0}$. The sufficient conditions ($c_m \geq \nu_m^* b_{m+1}$ for all m) given there for $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ to be isometric to an interval are also sufficient conditions for $\dot{A}_{L_0,0} = \dot{N}_{L_0,0}$.

Suppose $\dot{A}_{L_0,0} \neq \dot{N}_{L_0,0}$. Then there exists a 0-minimal curve $\gamma : \mathbb{R} \to \mathbb{T}^n$ such that $[\alpha(\gamma)] \neq [\omega(\gamma)]$. The Lagrangian L_0 has the form discussed in §6. The Proposition in §6 gives a formula for $\rho_{L_0,0}$. Since γ is 0-minimal, it achieves the infimum in that formula. However, the proof of the Proposition in §5 shows that this infimum is not achieved. This contradiction shows that there is no such γ. Consequently, $\dot{A}_{L_0,0} = \dot{N}_{L_0,0}$.

So far, we have obtained an L_0 such that $\bar{A}_{L_0,0}$ is isometric to an interval and $\dot{A}_{L_0,0} = \dot{N}_{L_0,0}$. In the remainder of this section, we obtain an example where $\bar{A}_{L_0,0}$ is isometric to an interval and $\dot{A}_{L_0,0} \neq \dot{N}_{L_0,0}$.

To do this, we first sharpen the results of the Proposition in §5 a little. We let R be an ultimate tooth of C^m and let R^* be the parallelepiped associated to R in §8 (i.e. the smallest parallelepiped that contains $C_R^{(m)}$). We let x and y be the bottom endpoints of the sides of R. We set

$$e_m = \inf \int_\gamma \sqrt{U}\, ds$$

where the infimum is taken over all rectifiable curves γ connecting x to y in $R^* \setminus \mathrm{EDZ}$. Here $\mathrm{EDZ} = \mathrm{EDZ}(R^*)$ is the enlarged deficit zone of R^* defined in §10. It is easily seen that $b_m \geq e_m \geq c_m$. (The last inequality was proved in §10.)

The proof of the Proposition in §5 is easily modified to show that $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is isometric to an interval if and only if $b_m^{(\infty)} > 0$ and $e_m \geq \nu_m^* e_{m+1}$ for all m. The argument that we have given in this section shows that $\dot{A}_{L_0,0} = \dot{N}_{L_0,0}$ if and only if $e_m > \nu_m^* e_{m+1}$ for all m. The construction at the end of §5 shows that it is possible to choose the numbers A_m so that U is of class $C^{2d-3,1-\epsilon}$ and $e_m \geq \nu_m^* e_{m+1}$ for all m, and $e_m = \nu_m^* e_{m+1}$ for at least one m. (In fact, it is possible to choose the numbers A_m so that $e_m = \nu_m^* e_{m+1}$ for all m.)

We may define $e_m^{(\infty)}$ in the same way that we defined $b_m^{(\infty)}$ and c_m^∞ in §5, viz.

$$e_m^{(\infty)} = \inf(e_m, \nu_m^* e_{m+1}, \nu_m^* \nu_{m+1} e_{m+2}, \dots).$$

Definition 2. For $x, y \in Z$, we set

$$\rho^{**}(x, y) = 2^{3/2} \sum_G e_{m(G)}^{(\infty)},$$

where the sum is taken over all primary gaps G in $Z_{x,y}^*$.

It is easy to see that, when $c_m \geq \nu_m^* b_{m+1}$ for all m, we have $b_m^{(\infty)} = e_m^{(\infty)} = c_m^{(\infty)}$ for all m, so $\rho^{**} = \rho^*$ (see §7). We do not know whether this is true under the weaker hypothesis that $e_m \geq \nu_m^* e_{m+1}$; however, this question is unimportant, since we may modify the Proposition in §7, as follows.

PROPOSITION. *Let $x, y \in Z$. In the case that $e_m \geq \nu_m^* e_{m+1}$ for all m, we have $\rho_{L_0,0}(x, y) = \rho^{**}(x, y)$.*

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

This modified Proposition may be proved in the same way as the Proposition in §7 and may be used to prove the generalization we have stated in this section of the Proposition in §5 (in the same way as the Proposition in §7 was used to prove the Proposition in §5).

18. *The failure of upper semi-continuity of the Aubry set*

In this section, we construct a modification L_1 of L_0 and show that the set-valued function $L \mapsto A_{L,c}$ is not upper semi-continuous at L_1.

We choose the numbers A_m that appear in the definition of U (§5) so that $b_m^{(\infty)} > 0$ for all m. Then, according to the Proposition in §5, $(\bar{A}_{L_0,0}, \rho_{L_0,0})$ is homeomorphic to an interval. (For what we do in this section, it is not necessary that it be isometric to an interval.)

We choose $x, y \in A_{L_0,0}$ such that $\rho_{L_0,0}(x, y) > 0$. We let γ be a smooth embedded curve in $\mathbb{T}^d$ that connects x to y and meets $A_{L_0,0}$ only in x and y. We let α^0 be a smooth 1-form on γ that vanishes near the endpoints of γ, is non-positive with respect to the given orientation of γ (from x to y), and satisfies

$$\int_\gamma \alpha^0 + \sqrt{2U}\, ds < 0. \tag{$\star$}$$

In addition, we suppose that there exists a number $\Delta > 0$ such that if γ' is a sub-arc of γ with endpoints ξ and η, then

$$\int_{\gamma'} \alpha^0 + \int_{\gamma'} \sqrt{2U}\, ds + \frac{1}{2}\rho_{L_0,0}(\xi, \eta) \geq \Delta \quad \text{and} \quad \int_{\gamma'} \alpha^0 + 2\int_{\gamma'} \sqrt{2U}\, ds \geq 0. \tag{$\dagger$}$$

Here γ' inherits its orientation from γ.

It is easy to see that such an α^0 exists if the fact that $\rho_{L_0,0}(\xi, \eta)$ is uniformly bounded below by a positive quantity for $\xi, \eta \in \gamma$ is taken into account. This fact follows from the facts that $\rho_{L_0,0}(x, y) > 0$ and γ intersects $A_{L_0,0}$ only in the endpoints of γ.

Next, we extend α^0 to a 1-form α on $\mathbb{T}^d$, as follows.

We choose a small tube T about γ in $\mathbb{T}^d$. Such a tube is diffeomorphic to $\gamma \times B^{n-1}$, where $B^{n-1} \subset \mathbb{R}^{n-1}$ is an n-ball centered at 0. We suppose that the diffeomorphism $T \approx \gamma \times B^{n-1}$ is chosen so that γ corresponds to $\gamma \times 0$. We let ρ be a C^∞ bump function on B^{n-1}, so $\rho(0) = 1$, $0 \leq \rho < 1$, and the support of ρ is in the interior of B^{n-1}. We let $\pi_1 : T \to \gamma$ and $\pi_2 : T \to B^{n-1}$ denote the projections on the respective factors and let

$$\alpha = \begin{cases} (\rho \circ \pi_2)(\pi_1^* \alpha^0) & \text{in } T, \\ 0 & \text{elsewhere.} \end{cases}$$

Since α^0 vanishes near the endpoints of γ, it is clear that α is a smooth 1-form on $\mathbb{T}^d$. Moreover, the support of α does not intersect $A_{L_0,0}$, provided that T is chosen small enough.

We express α in the form

$$\alpha = V_1(\theta)\, d\theta_1 + \cdots + V_d(\theta)\, d\theta_d$$

and set

$$L_1(\theta, \dot{\theta}) = L_0(\theta, \dot{\theta}) + V_1(\theta)\dot{\theta}_1 + \cdots + V_d(\theta)\dot{\theta}_d.$$

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

It follows from the conditions (†) that $A_{L_1,0} = A_{L_0,0}$ if T is chosen small enough.

We set $U^N = \sum_{m=0}^{N} A_m u_m$, where the A_m are the positive numbers that appear in the definition of U (§5). Thus, $U^N \to U$ as $N \to \infty$. We set $L_0^N = \frac{1}{2}\sum \dot{\theta}_i^2 + U^N(\theta_1, \ldots, \theta_d)$ and $L_1^N = L_0^N + V_1 \cdot \dot{\theta}_1 + \cdots + V_d \cdot \dot{\theta}_d$. Thus, $L_i^N \to L_i$ (for $i = 0, 1$) as $N \to \infty$.

By the definition of U^N, it vanishes identically on a neighborhood of $Z = Z_U$. It follows that

$$\inf_{\beta} \int_a^b L_1^N(\beta(t), \beta(t))\, dt \leq 0,$$

where the infimum is taken over all absolutely continuous curves $\beta : [a, b] \to \mathbb{T}^d$ such that $\beta(a) = y$ and $\beta(b) = x$.

Moreover,

$$\begin{aligned}\lim_{N\to\infty} \int L_1^N(\gamma(t), \dot{\gamma}(t))\, dt &= \int L(\gamma(t), \dot{\gamma}(t))\, dt \\ &= \int_\gamma \alpha^0 + \int_\gamma \sqrt{2U}\, ds < 0,\end{aligned}$$

where the integrations are taken over the domain of γ. Hence

$$\int L_1^N(\gamma(t), \dot{\gamma}(t))\, dt < 0,$$

for large N. Since γ connects x to y, the concatenation $\gamma * \beta$ is a closed curve. For N large enough and β suitably chosen the action of $\gamma * \beta$ is negative.

Let μ_N be a 0-minimal invariant probability measure for L_1^N. The argument above, together with Tonelli's Theorem and the Krylov–Bogoliuboff argument, shows that

$$\lim_{N\to\infty} \sup A(\mu_N) \leq \int_\gamma \alpha^0 + \int_\gamma \sqrt{2U}\, ds < 0,$$

where $A(\mu_N)$ denotes the average action of μ_N with respect to the Lagrangian L_1^N, i.e.

$$A(\mu_N) = \int L_1^N\, d\mu_N.$$

Let $\pi : T\mathbb{T}^d \to \mathbb{T}^d$ denote the projection. If $\pi(\operatorname{supp} \mu_N) \cap \operatorname{supp} \alpha = \phi$, then $\int L_1^N d\mu_N = \int L_0^N d\mu_N$. However, the latter is non-negative, since $L_0^N = \frac{1}{2}\sum d\dot{\theta}_i^2 + U_N$ and $U_N \geq 0$. Since $\int L_1^N d\mu_N < 0$ for large N, this gives a contradiction. We conclude that $\pi(\operatorname{supp} \mu_N) \cap \operatorname{supp} \alpha \neq \phi$ for large N. Hence $A_{L_1^N,0} \cap \operatorname{supp} \alpha \neq \phi$, for large N. Hence

$$\limsup A_{L_1^N,0} \cap \operatorname{supp} \alpha \neq \phi.$$

On the other hand, we saw above that $A_{L_1,0} = A_{L_0,0}$. Moreover, α was constructed so that $A_{L_0,0} \cap \operatorname{supp} \alpha = \phi$. Hence

$$\limsup A_{L_1^N,0} \not\subset A_{L_1,0}.$$

This is the failure of upper semi-continuity that we sought.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

19. *Total disconnectedness of $A_{L_0,0}$ when U is of class C^{2d-2}*

It is not possible to improve the differentiability class of U in Theorem 1. Suppose that U is C^{2d-2}. It follows from the Proposition in §6 that $\rho_{L_0,0}(x, y) \leq C\|x - y\|^d$ for a suitable constant C and for all $x, y \in A_{L_0,0}$. The total disconnectedness of $A_{L_0,0}$ then is an immediate consequence of Proposition 1 of [**Mat3**].

20. *Open problems*

It is known [**HW**] that a compact metric space has vanishing topological dimension if and only if it is totally disconnected. Thus, the problem posed in §2 can be reformulated to ask whether $\bar{A}_{L,c}$ has vanishing topological dimension when L is C^∞.

A stronger condition than vanishing topological dimension is vanishing box dimension. It is easily seen that if L_0 is C^∞ and has the form given by ($\star$) in §2, then $\bar{A}_{L_0,0}$ has vanishing box dimension. This suggests that the following problem might have a positive answer.

Problem. If L is C^∞, does the box dimension of $\bar{A}_{L,c}$ vanish?

Obviously, a positive solution of this problem would imply a positive solution of the problem in §2.

Our studies on Arnold diffusion (not yet written) suggest that it is useful to know that $\bar{A}_{L,c}$ is 'small' in some sense of dimension (e.g. it has vanishing topological or box dimension). However, to prove that Arnold diffusion is a generic phenomenon, it should be enough to show that $\bar{A}_{L,c}$ is generically 'small' in an appropriate sense. This suggests the following problem.

Problem. Let $\mathcal{S}^r_{\text{top}}$ (respectively $\mathcal{S}^r_{\text{box}}$) be the set of all C^r Lagrangians such that $\bar{A}_{L,c}$ has vanishing topological (respectively box) dimension for all $c \in H^1(M, \mathbb{R})$. Is $\mathcal{S}^r_{\text{top}}$ (respectively $\mathcal{S}^r_{\text{box}}$) dense in the set of all C^r Lagrangians with the C^r topology? Residual?

In another direction, we have the following.

Problem. Is the set-valued function $c \mapsto A_{L,c}$ upper semi-continuous?

Acknowledgements. The author wishes to thank Eileen Olszewski for the care she took typing this manuscript and Albert Fathi for his encouragement and patience. Supported by NSF grant DMS-0245336.

References

[A] P. Assouad. Plongements lipschitziens dans $\mathbb{R}^n$. *Bull. Soc. Math. France* **111** (1983), 429–448.

[B] V. Bangert. Mather sets for twist maps and geodesics on tori. *Dynamics Reported*. Vol. 1. Wiley, Chichester, 1988, pp. 1–56.

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

[C] M. J. D. Carneiro. On minimizing measures of the action of autonomous Lagrangians. *Nonlinearity* **89** (1995), 1077–1085.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press

[F] A. Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math*. **327** (1998), 267–270.

[FaM] A. Fathi and J. N. Mather. Failure of convergence of the Lax–Oleinik semi-group in the time-periodic case. *Bull. Soc. Math. France* **128** (2000), 473–483.

[FoM] 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.

[HW] W. Hurewicz and H. Wallman. *Dimension Theory (Princeton Mathematical Series, 4)*. Princeton University Press, Princeton, NJ, 1941.

[Mat1] J. N. Mather. Action minimizing measures for positive definite Lagrangian systems. *Math. Z.* **207** (1991), 169–207.

[Mat2] J. N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)* **43** (1993), 1349–1386.

[Mat3] J. N. Mather. A property of compact, connected, laminated subsets of manifolds. *Ergod. Th. & Dynam. Sys.* **22** (2002), 1507–1520.

[Mat4] J. N. Mather. Total disconnectedness of the quotient Aubry set in low dimensions. *Comm. Pure Appl. Math*. **56** (2003), 1178–1183.

[Mat5] J. N. Mather. Arnold diffusion, I: announcement of results. *Contemporary Mathematics and its Applications, Fundamental Directions* April–May 2003 (Russian transl.), original to appear in *J. Math. Sci*.

https://doi.org/10.1017/S0143385704000446 Published online by Cambridge University Press