Ergod. Th. & Dynam. Sys. (1997), **17**, 187–203
Printed in the United Kingdom

Shadowing property of geodesics in Hedlund's metric

MARK LEVI
Math. Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

(*Received* 29 *July* 1994)

Abstract. In this paper we show that the geodesic flow in a Hedlund-type metric on the 3-torus possesses the shadowing property. This implies, in particular, that any rotation vector is represented by a geodesic, a fact that in the two-dimensional case is given by the Aubry–Mather theory, while in the higher-dimensional case is still unknown.

1. *Introduction*

The Aubry–Mather theory [**A**, **MA1**] guarantees the existence of invariant sets in area-preserving annulus maps satisfying some mild monotone twist assumptions. These sets can be thought of as remnants of invariant KAM circles. Remarkably, every rotation number is represented by one such Mather set. An essentially equivalent Riemannian geometric counterpart of an area-preserving annulus map is the geodesic flow on the 2-torus—more precisely, the two problems are applications of one more general theory [**B2**]. The Mather sets for annulus maps correspond to minimal geodesics on the covering plane of the 2-torus (by the definition of a minimal geodesic it realizes the shortest path between any of its two points on the covering plane). A crucial property of Mather sets is the fact that they are global minimizers of the action [**B1**, **BK**, **BP**, **G**, **H**, **K**, **M**, **M2**, **MA1**, **MA2**, **MO1**, **MO2**, **MO3**]; it is therefore natural to attempt to extend the Aubry–Mather theory to higher-dimensional symplectic maps or flows by seeking action-minimizing orbits. Such an attempt cannot succeed, as was pointed out by Hedlund [**H**], who produced a Riemannian metric on a 3-torus in which only three rotation vectors are represented by minimal geodesics. More recently, Bangert [**B1**] gave a detailed study of the existence and properties of minimal geodesics and showed that there exist restrictions on the rotation vectors of arbitrary minimal geodesics.

Hedlund's idea of non-existence of many rotation vectors leads to the question which seems to be a natural route for generalization of the Aubry–Mather theory: for a general Riemannian metric on the 3-torus L, does there exist a locally minimal geodesic (i.e. minimal among the geodesics in a tubular neighborhood) for any rotation vector? This question still remains open; the goal of this paper is to give a solution in a particular case of Hedlund's metric. In dimension 2 this problem has been solved by Hedlund [**H**] for an arbitrary Riemannian metric.

2. *Definitions and the results*

Following Hedlund, we define a conformal Riemannian metric $dh = \rho(\mathbf{x})ds$ on the 3-torus

$$\mathbf{T}^3 = \{(x_1, x_2, x_3) \bmod 1\},$$

where $ds^2 = dx_1^2 + dx_2^2 + dx_3^2$ is the Euclidean metric and where $\rho(\mathbf{x})$ is small in certain 'tubes' and $\rho = 1$ elsewhere. To be precise, we define the lines $L_1,\ L_2,\ L_3$ (which will be the axes of the 'tubes') as

$$\begin{aligned} L_1 &= \{t\mathbf{e}_1,\ t \in \mathbb{R}\} \\ L_2 &= \{t\mathbf{e}_2 + \tfrac{1}{2}\mathbf{e}_3,\ t \in \mathbb{R}\} \\ L_3 &= \{t\mathbf{e}_3 + \tfrac{1}{2}\mathbf{e}_1 + \tfrac{1}{2}\mathbf{e}_2,\ t \in \mathbb{R}\}. \end{aligned}$$

Furthermore, for any $\epsilon > 0$ we define the tube $T_i(\epsilon)$ as follows (Figure 1):

$$T_i(\epsilon) = \{\mathbf{x} \in \mathbb{R} : \text{dist}(\mathbf{x}, L_i) < \epsilon\}.$$

All the tubes are disjoint if $0 < \epsilon < \frac{1}{4}$. Together with $T_i(\epsilon)$ we consider all their integer translates $T_i(\epsilon) + N,\ N \in \mathbb{N}$ in $\mathbb{R}^3$, and define the Hedlund metric as follows.

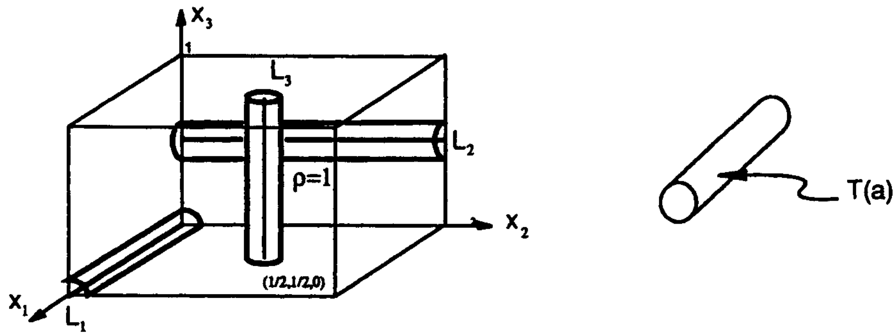

FIGURE 1. Definition of the Hedlund-type metric.

Definition 1. The *Hedlund metric* is defined by a function $\rho \in C^0(\mathbb{R}^3)$ of period 1 in each coordinate and given by

$$\rho_\epsilon(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \text{ is outside } T_i(\epsilon) + \mathbb{Z},\ i = 1, 2, 3; \\ (\epsilon^2 + \epsilon^{-2}(1 - \epsilon^2)r^2)^{1/2}, & \text{if } \mathbf{x} \in T_i(\epsilon) + \mathbb{Z},\ i = 1, 2, 3. \end{cases}$$

(Here, r is the distance from $\mathbf{x}$ to the nearest line L_i.)

The choice of ϵ will be specified in the next section.

Before stating the main theorem, we define the pseudogeodesics and give their symbolic encoding. The following definition of the pseudogeodesics is motivated by the intuition that there exist geodesics which 'ride' along the tube (approximately) for a prescribed distance and then 'switch' to another prescribed tube at the moment of closest passage, with the ride-and-switch sequence prescribed arbitrarily (in a consistent way) for the entire infinite length of that geodesic.

Definition 2. A *pseudogeodesic* is a continuous broken line consisting of a bi-infinite sequence of abutting mutually orthogonal rectilinear segments of two alternating types, which one can call 'riding' and 'switching' (Figure 2). The 'riding' interval is, by the definition, any segment of a line $L_i + N$, $N \in \mathbb{N}$, whose endpoints have half-integer or integer coordinates x_i. The 'switching' segment is any segment of length $\frac{1}{2}$ which connects two lines $L_i + N$, $L_j + M$, $i \neq j$, $N, M \in \mathbb{N}$ which are at the distance $\frac{1}{2}$.

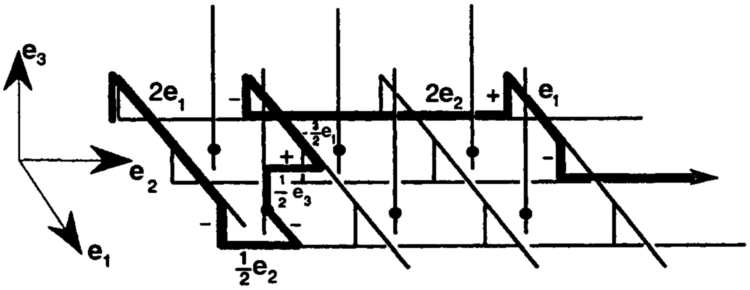

FIGURE 2. The local pseudogeodesic γ_σ^0.

Symbolic encoding of the pseudogeodesics. To specify the pseudogeodesic completely (modulo an integer translation), we have to specify:

- the'ride': $l \cdot \mathbf{e}_i$, $l \in \frac{1}{2}\mathbb{N}$, $i \in \{1, 2, 3\}$;
- the sign of the direction of the switch: $s \in \{+, -\}$.

Note that the direction itself is given by the common perpendicular to the two rides just before and just after the switch.

Thus, the set of all pseudogeodesics (identified modulo an integer translation in $\mathbb{R}^3$) is in one-to-one correspondence with the space of all sequences of the form

$$\sigma = \ldots (l_{-1}\mathbf{e}_{i_{-1}}, s_{-1})(l_0\mathbf{e}_{i_0}, s_0)(l_1\mathbf{e}_{i_1}, s_1) \ldots,$$

where

$$\mathbf{e}_{i_n} \neq \mathbf{e}_{i_{n+1}} \tag{1}$$

and

$$2l_i \text{ is even if } \mathbf{e}_{i_{n-1}} = \mathbf{e}_{i_{n+1}} \text{ and odd otherwise.} \tag{2}$$

As an example, a piece of a pseudogeodesic given by

$$\ldots (2\mathbf{e}_1, -)(\tfrac{1}{2}\mathbf{e}_2, -)(\tfrac{1}{2}\mathbf{e}_3, +)(-\tfrac{3}{2}\mathbf{e}_1, -)(2\mathbf{e}_2, +)(\mathbf{e}_1, -) \ldots$$

is shown in figure 2.

To state the main results we define the tubular neighborhood $\Gamma_\sigma^\varepsilon$ of the pseudogeodesic γ_σ^0 and the patches $Q(p)$.

Definition 3. The patches. Let q be an intersection point of γ_σ^0 with the boundary $\partial T(\varepsilon)$ of one of the tubes $T(\varepsilon) = T_i(\varepsilon) + N$, $N \in \mathbb{N}$. We define the patch $Q(q)$ as a rectangular neighborhood on ∂T of the point q, given by

$$Q(q) = \left\{ \mathbf{x} \in \partial T(\varepsilon) : |s_1| \leq \bar{s}_1 \equiv \sqrt{\varepsilon},\ |s_2| = \bar{s}_2 \leq \frac{\pi}{4}\varepsilon \right\},$$

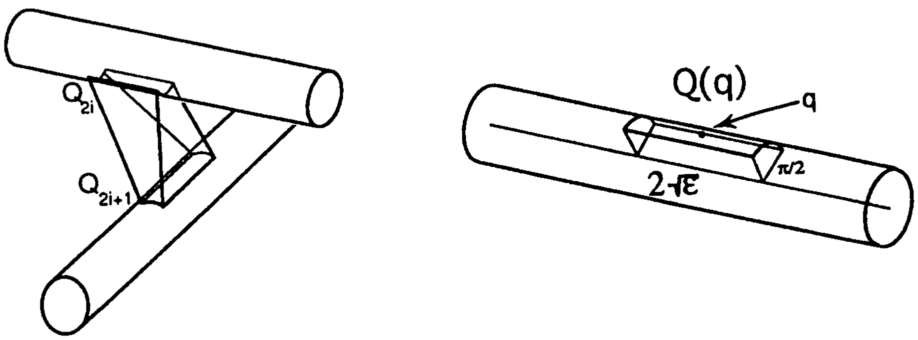

FIGURE 3. Definition of patches $Q(q)$, $\Gamma_\sigma^\varepsilon$.

where s_1 measures the displacement from q along the axial direction of the cylinder $\partial T(\varepsilon)$, while s_2 measures the displacement along the meridian of the cylinder.

Definition 4. The neighborhood $\Gamma_\sigma^\varepsilon$ of γ_σ^0. Given a pseudogeodesic γ_σ^0, let $\{q_i\}_{i=-\infty}^{\infty}$ be the sequence of the intersection points of γ_σ^0 with the boundaries of the tubes $T_i(\varepsilon)+N$, $N \in \mathbb{N}$, and let $Q_i = Q(q_i)$ be the patches as in the previous definition. To be specific, let the exit patches Q_{2i} be even. We define the *transition box*

$$B_{2i} = \text{convex hull of } \{Q_{2i} \cup Q_{2i+1}\}.$$

For the duration of this sentence only, denote by x_i the longitudinal coordinate in the one tube T^i on which both patches Q_{2i-1}, Q_{2i} lie, and define (see Figure 3)

$$\Gamma_\sigma^\varepsilon = \left(\bigcup_{i\in\mathbb{Z}} B_{2i}\right) \cup \left(\bigcup_{i\in\mathbb{Z}} \{\mathbf{x} : \mathbf{x} \in T^i,\ \inf x_i(Q_{2i-1}) \le x_i(\mathbf{x}) \le \sup x_i(Q_{2i})\}\right).$$

THEOREM 1. (Shadowing) *For any sequence σ satisfying the above admissibility conditions (1) and (2) with the corresponding pseudogeodesic γ_σ^0, there exists a unique geodesic γ_σ which belongs to the neighborhood $\Gamma_\sigma^\varepsilon$ of γ_σ^0.*

THEOREM 2. (Hyperbolicity) *If two geodesics γ_{σ^1}, γ_{σ^2} whose existence is given in Theorem 1 have eventually coinciding itineraries: $\sigma_n^1 = \sigma_n^2$ for all n large enough, then γ_{σ^1} and γ_{σ^2} approach each other exponentially in the future. A similar statement holds for the past.*

THEOREM 3. *Given any two geodesics γ_{σ^1} and γ_{σ^2}, there exists a heteroclinic geodesic $\gamma_{\sigma^{12}}$ asymptotic to γ_{σ^1} in the past and to γ_{σ^2} in the future. In particular, any pair of periodic geodesics γ_{σ^1} and γ_{σ^2} is heteroclinically connected. In other words, the unstable manifold of any periodic geodesic intersects the stable manifolds of all other periodic geodesics.*

Before stating some corollaries we define the *rotation vector* of a geodesic as the limit (if it exists)

$$\mathbf{r} = \lim_{t\to\infty} \frac{\gamma(t)}{\|\gamma(t)\|},$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^3$. We also define the *average action* as

$$\alpha = \lim_{t\to\infty} \frac{|\gamma([0,t])|_{H(\epsilon)}}{|\gamma([0,t])|_E} \equiv \lim_{S\to\infty} \frac{1}{S}\int_0^S \rho(\gamma)\,ds.$$

These definitions are suggested by the analogy with the variational problem

$$\delta \int_{-\infty}^{\infty} L(x,\dot{x},t)\,dt = 0$$

corresponding to the Euler–Lagrange equations

$$\frac{d}{dt}L_{\dot{x}} - L_x = 0.$$

The above theorems imply the following corollaries.

COROLLARY 1. *There exist uncountably many geodesics with any given (unit) rotation vector. Moreover, the rotation vector can be prescribed for the past and the future independently.*

COROLLARY 2. *There exist geodesics which visit every rotation vector arbitrarily closely. More precisely, there exists a geodesic γ such that for any unit vector $\mathbf{u}$ there exists a subsequence $\{t_n\} \to \infty$ such that*

$$\lim_{n\to\infty} \frac{\gamma(t_n)}{|\gamma(t_n)|} = \mathbf{u}.$$

Equivalently, the geodesic projected onto the unit sphere in the covering space $\mathbb{R}^3$ is dense.

COROLLARY 3. *Given any number $\alpha \in [\epsilon, \frac{1}{2}]$ and any rotation vector $\mathbf{u}$, there exists a geodesic with $\mathbf{u}$ as the rotation vector and with α as the average action:*

$$\alpha = \lim_{S\to\infty} \frac{1}{S}\int_0^S \rho(\gamma)\,ds \quad \textit{and} \quad \mathbf{u} = \lim_{t\to\infty} \frac{\gamma(t)}{|\gamma(t)|},$$

where S is the Euclidean arclength. Thus, the average Hedlund length per unit Euclidean length can be prescribed arbitrarily for any given rotation vector.

COROLLARY 4. *The geodesic flow on $\mathbf{T}^3$ with Hedlund's metric has positive topological entropy.*

Mechanical and optical equivalents. We use the well-known equivalence of the Riemannian geometry with mechanics expressed by the Maupertuis principle to observe that the geodesics we are describing are also the orbits of the particle moving with total energy $E = 1$ in the potential given by $V(\mathbf{x}) = 1 - \rho^2(\mathbf{x}) > 0$ inside the tubes and by $V = 0$ outside. These trajectories satisfy

$$\ddot{\gamma} = -\nabla V(\gamma)$$

with

$$\tfrac{1}{2}\dot{\gamma}^2 + V(\gamma) = 1.$$

Inside the Hedlund tubes the potential $V > 0$ is higher than $V = 0$ outside, and thus the sliding motions along the tubes are precariously balanced; in fact, we show that the behavior described here is hyperbolic. It can also be pointed out that the repulsion from the tubes is similar to the repulsion from the potential peaks in the Frenkel–Kontorova model [**FK**] which was used by Aubry [**A**] to show the non-existence of invariant circles in annulus maps far from the KAM case. A crucial property distinguishing the Hedlund problem considered here from the lower-dimensional case of the Frenkel–Kontorova model is that the potential well, which is $\mathbb{R}^3$ with the tubes cut out, is connected.

This problem admits an optical interpretation as well: the geodesics are the rays in the medium with the index of refraction $\rho(\mathbf{x})$, i.e. with the speed of light given by $c(\mathbf{x}) = 1/\rho(\mathbf{x})$. In this interpretation the tubes defocus the rays: a beam sent along the tube tends to spread; this spreading is responsible for the shadowing property.

The main theorem stated above can be viewed as a singular perturbation result for the Riemannian metric on $\mathbb{T}^3$. We also note that the tubular neighborhoods of pseudogeodesics constructed below give the isolating blocks (Conley [**C**]) for the curvature flow; the fixed points of this flow are the geodesics.

3. *Outline of the proof of the existence and hyperbolicity of shadowing geodesics*

3.1. *Basic lemmas.* Let σ be an arbitrary (admissible) sequence with the corresponding pseudogeodesic γ_σ^0. Let $\dots Q_0, Q_1, \dots$ be the sequence of patches constructed as above. The following lemmas contain key geometric ingredients in the proof. In what follows, x denotes the longitudinal coordinate in the tube T.

LEMMA A. (The connection lemma) *Given any pair of points p, $q \in \partial T(\varepsilon)$ with $|x(p) - x(q)| \geq \frac{1}{8}$, where $T(\varepsilon) = T_i(\varepsilon)$ is any of the tubes defined above with x denoting the longitudinal coordinate in T, there exists a unique geodesic segment in the Hedlund metric with p and q as its endpoints and lying entirely inside $T(\varepsilon)$.*

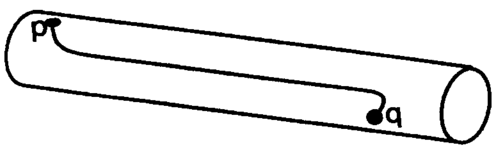

FIGURE 4. The unique geodesic in T connecting two points on $p, q \in \partial T$.

LEMMA B. *Let $p, q \in \partial T$ and let (s_1, s_2) and (s_3, s_4) be the coordinates of p and q on ∂T, where the longitudinal coordinates are s_1, s_3. In these coordinates the Hedlund length $L^\varepsilon(s)$, $s = (s_1, s_2, s_3, s_4)$ of the geodesic segment from Lemma A satisfies the estimates*

$$|DL^\varepsilon|, \quad |D^2L^\varepsilon| = O(\varepsilon), \tag{3}$$

uniformly in p, q for all $|s_1 - s_3| \geq \frac{1}{8}$.

For the next lemma, recall that Q_{2i} are chosen to be the exit sections. Without loss of generality let $i = 0$, so that Q_0 and Q_1 are the consecutive exit and entrance patches respectively.

LEMMA C. *Let $L^1 : Q_0 \times Q_1 \to \mathbb{R}$ be the Euclidean distance between $q_0 \in Q_0$ and $q_2 \in Q_1$ expressed in the coordinates (s_1, s_2) and (s_3, s_4) on the patches Q_0 and Q_1: $L^1 = L^1(s_1, s_2, s_3, s_4)$. Then*

$$D^2 L^1 \geq cI, \quad c > 0, \ \forall q \in Q = Q_0 \times Q_1,$$

with c independent of ε for all $0 < \varepsilon \leq \frac{1}{8}$.

LEMMA D. *The flow φ^t of the gradient vector field $\dot{s} = -\nabla L^1(s)$ on $Q = Q_0 \times Q_1$ (with s as in Lemma C) takes Q into itself for $t \geq 0$; moreover, for any supporting unit* outward *vector $n(q)$ on ∂Q we have*

$$D_n L^1 = n \cdot \nabla L^1 = -\dot{s} \cdot n \geq \tfrac{1}{2}\sqrt{\varepsilon} \ \text{for all } q \in \partial Q. \tag{4}$$

The proofs of these lemmas are given in §3.3.

3.2. *Proof of Theorems 1–3.* In this section we use Lemmas A–D; their proof is given in the next section.

Proof of Theorem 1. Picking any admissible sequence σ with the corresponding pseudogeodesic γ_σ^0, let $\dots g_{-1} g_0 g_1 \dots$ be the sequence of intersections of γ_σ^0 with the boundaries of the tubes. Let Q_i be the patch centered at g_i, cf. Definition 3. We first show (in step 1) that there exists a geodesic of finite length of any given itinerary and then prove the existence of the infinite geodesic (step 2).

Step1. We show that there exists a unique geodesic segment in Γ_σ^0 connecting any pair of points $q_m \in Q_m$, $q_n \in Q_n$. Without loss of generality we fix $q_1 \in Q_0$ and $q_{2N} \in Q_{2N}$, make an arbitrary choice of intermediate points $q_i \in Q_i$, $2 \leq i \leq 2N-1$ and consider the broken geodesic $q_1 q_2 \dots q_{2N}$; it is uniquely defined by q_i according to Lemma A. The length of this broken geodesic is a function of coordinates (s_1^i, s_2^i) of q_i (cf. Definition 3); for convenience we will identify the point q_i with the pair of its coordinates on the patch: $q_i \equiv (s_1^i, s_2^i)$. Thus, the length function is defined on a box in $\mathbb{R}^{4(N-1)}$. Before proceeding further it is convenient to treat each Euclidean segment of the geodesic as one object, i.e. to define

$$p_k = \begin{pmatrix} q_{2k} \\ q_{2k+1} \end{pmatrix} \in B_k \subset \mathbb{R}^4, \quad k = 1, \dots, N-1;$$

here $B_k = Q_{2k} \times Q_{2k+1}$ is a box in $\mathbb{R}^4$; we have also assumed that the even points q_{2k} are the exit points from the tubes (Figure 5). We define p_0 and p_{2N} as

$$p_0 = \begin{pmatrix} 0 \\ q_1 \end{pmatrix} \quad \text{and} \quad p_N = \begin{pmatrix} q_{2N} \\ 0 \end{pmatrix};$$

these vectors are fixed since q_1 and q_{2N} have been fixed. The length of our broken geodesic $L: B_1 \times \dots \times B_{N-1} \to \mathbb{R}$ takes the form

$$L(p_1, \dots, p_{N-1}) = L^\epsilon(p_1) + L^1(p_1) + L^\epsilon(p_1 p_2) + \dots + L^1(p_{N-1}) + L^\epsilon(p_{N-1}).$$

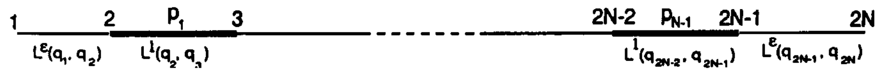

FIGURE 5. The length function consists of the Euclidean (bold segments) and the 'Hedlund' portions (thin segments).

We show that L has a unique minimum and no other extrema in the box $B = B_1 \times \cdots \times B_{N-1}$, thus proving the existence of a unique geodesic segment $q_0 q_{2N}$. First we note that the gradient flow of $\dot{p} = -\nabla L(p)$, $p = (p_1, \ldots, p_{N-1}) \in B$ takes B into itself, as follows from Lemmas D and B; indeed, pick $p \in \partial B$, and show that $-\nabla L(p)$ points inwards. To that end, let $n = (n_1, \ldots, n_{N-1})$ be the unit normal to the face of B at p; it suffices to consider the case when p lies on the face of the highest dimension, say, when $n = (n_1, 0, \ldots, 0)$. Then

$$\dot{p} \cdot n = -\left(\frac{\partial L^{\epsilon}}{\partial p_1}(p_1) + \frac{\partial L^1}{\partial p_1}(p_1)\right) \cdot n_1 \le -\frac{1}{4}\sqrt{\epsilon} < 0.$$

This shows at once the existence of a minimum inside B. Now, at any point $p \in B$, the Hessian of L is close to the block-diagonal matrix

$$D^2 L(p) = \operatorname{diag}(D^2 L^1(p_1), D^2 L^1(p_{N-1})) + O(\epsilon),$$

with 4×4 blocks on the diagonal. The smallness of the terms arising from L^{ϵ} follows from Lemma B. We conclude by Lemma C that

$$D^2 L(p) \ge cI > 0,$$

so that any extremum in B is a local minimum and in fact the only extremum since B is topologically a ball. This proves the existence of a unique geodesic connecting q_1 and q_{2N} with the prescribed itinerary.

Step 2. The existence of an infinite geodesic is proven by a diagonal process. Pick any admissible sequence σ with the corresponding pseudogeodesic γ_{σ}^0 and the corresponding sequence of sections

$$\ldots Q_{-1}, Q_0, Q_1 \ldots$$

which γ_{σ}^0 crosses at $q_i \in Q_i$, $i \in \mathbb{Z}$. Consider an infinite curve γ_N obtained by replacing the finite segment of γ_{σ}^0 between q_{-N} and q_N by a geodesic in $\Gamma_{\sigma}^{\varepsilon}$ whose existence was proven in the previous step, while leaving the rest of γ_{σ}^0 intact. Letting $N \to \infty$, we extract a subsequence from $\{\gamma_N\}$ which converges uniformly between Q_{-1} and Q_1; denote this subsequence by

$$\gamma_1^1, \gamma_2^1, \ldots.$$

From this subsequence, in turn, we extract a subsequence converging between Q_{-2} and Q_2:

$$\gamma_1^2, \gamma_2^2, \ldots.$$

Proceeding inductively we obtain a sequence

$$\gamma_1^n, \ldots, \gamma_2^n, \ldots$$

for any n which converges between Q_{-n} and Q_n. It remains to observe that the diagonal sequence

$$\gamma_1^1, \gamma_2^2, \ldots$$

converges on *any* finite interval; its limit is the desired geodesic.

Step 3. We show that a geodesic segment in $\Gamma_\sigma^\varepsilon$ between $q_{-N} \in Q_{-N}$ and $q_N \in Q_N$ depends smoothly on its endpoints $q_{\pm N}$.

We have

$$L(q_{-N}, q_N) = L(q_{-N}, q_{-N+1}) + \cdots + L(q_{N-1}, q_N),$$

where $(q_{-N+1} \cdots q_{N-1})$ is determined by the requirement that it be a critical point. By the properties of L^1 and L^ε this point is in fact a non-degenerate minimum (by the block-diagonal nature of D^2L), and thus depends smoothly on q_{-N}, q_N. The global (in $Q_{-N+1} \times \cdots \times Q_{N-1}$) uniqueness of this point follows from positive invariance of $Q_{-N+1} \times \cdots \times Q_{N-1}$ under $-\nabla L$ (Lemma D) and from Lemma C. The proof of Theorem 1 is complete. □

Proof of Theorems 2 and 3. It suffices to show that any two geodesic segments γ^0, γ^1 in $\Gamma_\sigma^\varepsilon$ which share the same finite sequence of sections

$$Q_1, \ldots, Q_{2N},$$

which they cross at the points $q_1^0, \ldots, q_{2N}^0$ and at $q_1^1, \ldots, q_{2N}^1$ respectively, satisfy, for some $c_1, c_2 > 0$ independent of N,

$$|q_k^1 - q_k^0| < c_1 \varepsilon^{c_2 k} |q_1^1 - q_1^0| + c_1 \varepsilon^{c_2(2N-k)} |q_{2N}^1 - q_{2N}^0|, \quad 1 < k < 2N.$$

Introduce a smooth deformation q_1^t, $0 \le t \le 1$ from q_1^0 to q_1^1, similarly for q_{2N}^t, and consider the connecting geodesic with corresponding intermediate points $q_k^t \in Q_k$, $k = 2, \ldots, 2N-1$, which are unique by Step 3. Denoting $v_k = v_k^t = (d/dt)q_k^t$, it suffices to prove that for some c_1, c_2 independent of ε,

$$|v_k| < c_1 \varepsilon^{c_2 k} |v_1| + c_1 \varepsilon^{c_2(2N-k)} |v_{2N}|. \tag{5}$$

To prove (5) we show that v_k satisfy a simple difference equation, equation (8) below, where $u_n = \binom{v_{2n}}{v_{2n+1}}$.

The equilibrium condition

$$\frac{\partial L}{\partial p_k} = 0$$

becomes

$$DL^1(p_k) + D_2 L^\varepsilon(p_{k-1} p_k) + D_1 L^\varepsilon(p_k p_{k+1}) = 0, \quad k = 1, \ldots, N-1,$$

where $D_i L(p_1, p_2)$ denotes the derivative with respect to $p_i = (s_1^i, s_2^i)$. Differentiating by t and denoting

$$u_k = \frac{\partial}{\partial t} p_k = \frac{\partial}{\partial t} \begin{pmatrix} q_{2k} \\ q_{2k+1} \end{pmatrix} \equiv \begin{pmatrix} v_{2k} \\ v_{2k+1} \end{pmatrix},$$

we obtain the linearized equation

$$H_k u_k + E_k^- u_{k-1} + E_k^+ u_{k+1} = 0, \tag{6}$$

where

$$H_k = D^2L^1(p_k) + D_2^2L^\varepsilon(p_{k-1}p_k) + D_1^2L^\varepsilon(p_k p_{k+1}),$$
$$E_k^- = D_1D_2L^\varepsilon(p_{k-1}p_k)$$
$$E_k^+ = D_1D_2L^\varepsilon(p_k p_{k+1}).$$

with

$$u_0 = \begin{pmatrix} 0 \\ v_1 \end{pmatrix} \quad \text{and} \quad u_N = \begin{pmatrix} v_{2N} \\ 0 \end{pmatrix}. \tag{7}$$

Multiplying by H^{-1} we rewrite (6) in the simple form

$$u_n = \varepsilon_n^- u_{n-1} + \varepsilon_n^+ u_{n+1}, \quad 1 \le n \le N-1, \tag{8}$$

with the boundary conditions (7). The 4×4 matrices $\varepsilon_n^\pm$ satisfy, for some $c > 0$ independent of ε,

$$2\|\varepsilon_n^+\| \le c\varepsilon, \quad 2\|\varepsilon_n^-\| \le c\varepsilon, \tag{9}$$

as follows from Lemmas B and C.

LEMMA E. *Any difference equation (8) satisfying (9) and with* any *prescribed boundary conditions u_0 and u_N has a unique solution. This solution satisfies*

$$|u_n| \le |u_0|(c\varepsilon)^n + |u_N|(c\varepsilon)^{N-n}. \tag{10}$$

We use this lemma to conclude that

$$|v_{2n}| + |v_{2n+1}| \le |v_1|(c\varepsilon)^n + |v_{2N}|(c\varepsilon)^{N-n},$$

which proves (5). This completes the proof of Theorems 2 and 3. □

Proof of Lemma E. By linearity it suffices to consider the special case of $u_N = 0$. We prove both the existence and the desired estimate

$$|u_n| \le |u_0|(c\varepsilon)^n \tag{11}$$

iteratively by starting with

$$u^0 = (u_1^0, \ldots, u_{N-1}^0) = (0, \ldots, 0)$$

and defining the iteration $u^k \to u^{k+1} = (u_1^{k+1}, \ldots, u_{N-1}^{k+1}) = \varphi(u^k)$ by sweeping to the right row by row (see Figure 6): given the row u^k, we find u^{k+1} by running m from 1 to $N-1$, that is

$$\begin{cases} u_m^{k+1} = \varepsilon_m^- u_{m-1}^{k+1} + \varepsilon_m^+ u_{m+1}^k, \quad m = 1, \ldots, N-1 & (12) \\ u_0^{k+1} = u_0, \quad u_N^{k+1} = 0. & (13) \end{cases}$$

To show that each iterate $u^k = (u_0^k, \ldots, u_N^k)$ satisfies (11), we assume inductively that (11) holds for u^k and show that it then holds for u^{k+1}. Using (12), (13), (9) and the inductive assumption gives

$$|u_1^{k+1}| = |\varepsilon_1^- u_0^{k+1} + \varepsilon_1^+ u_2^k| \le \frac{c}{2}\varepsilon(|u_0| + |u_2^k|) \le \frac{c}{2}\varepsilon(1 + (c\varepsilon)^2)|u_0| < c\varepsilon|u_0|. \tag{14}$$

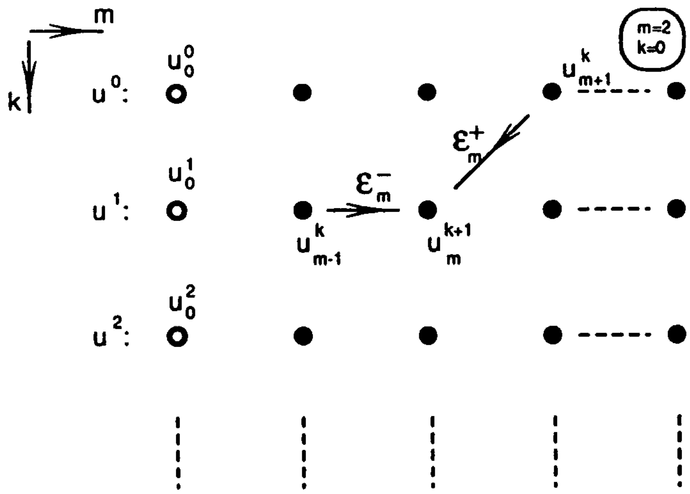

FIGURE 6. Proof of Lemma E: the 'sweeping' iteration.

Proceeding inductively from u_1^{k+1} to $u_2^{k+1}, \dots,$ to u_N^{k+1} identically with the last step shows that $u^{k+1} = \varphi(u^k)$ satisfies (11).

To prove that the iterates converge we define the 'energy'

$$E(u) = E(u_1, \dots, u_{N-1}) = \sum_{m=1}^{N-1} (u_m - \varepsilon_m^- u_{m-1} - \varepsilon_m^+ u_{m+1})^2, \tag{15}$$

which is a Lyapunov function for our iteration process φ:

$$E(\varphi(u)) \le E(u), \tag{16}$$

as can be easily checked. Moreover, we have the equivalence

$$\varphi(u) = u \Longleftrightarrow E(\varphi(u)) = E(u), \tag{17}$$

from the definition of φ and of E. Since φ maps the cube $\{(u_1, \dots, u_{N-1}) : |u_i| \le |u_0|\}$ into itself for ε small enough, the sequence of iterates $\varphi^k(u)$ has a convergent subsequence; let u^∞ be its limit. Since E is monotone along the sequence $\varphi^k(u)$ and positive, we obtain in the limit

$$E(\varphi(u^\infty)) = E(u^\infty),$$

i.e. $\varphi(u^\infty) = u^\infty$ by (17), proving that u^∞ is a solution. The exponential decay estimates survive the limiting process. Uniqueness follows from exponential estimates and linearity. □

3.3. *Proof of basic Lemmas C, D, A and B.*

Proof of Lemma C. As suggested by Figure 7, two of the eigenvalues of the Hessian of L^1 are $O(1/\varepsilon)$ which is large, much larger than the lemma states. This, in fact, is shown in the proof below but not expressed in the statement of the lemma since we are only concerned with the lower bounds on the Hessian.

Fix an arbitrary pair of points $q_i \in Q_i$, $i = 0, 1$, and let $\mathbf{t} = (t_1, t_2)$ and $\mathbf{s} = (s_1, s_2)$ be the coordinates on Q_i. By shifting the origin we assume that $\mathbf{s} = 0$, $\mathbf{t} = 0$ correspond to q_0 and q_1 (see Figure 7). Without loss of generality, we assume that t_1 and s_2 are the longitudinal coordinates along the axes of respective cylinders; we also assume that the cylinders are aligned along the x_2 and x_1 axes in $\mathbb{R}^3 = \{x_1, x_2, x_3\}$ respectively.

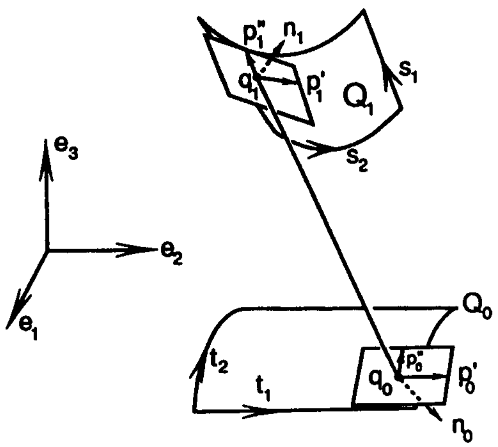

FIGURE 7. Proof of Lemma C.

In order to compute $D^2|q_1 - q_0|$ we let $\mathbf{n}_i$ be the unit normal to the surface Q_i of the ith cylinder at q_i, in the direction away from the other surface. Let $\mathbf{p}_0'$, $\mathbf{p}_0''$ be the unit vectors in the tangent planes $T_q Q_0$ corresponding to the coordinates $\mathbf{t} = (t_1, t_2)$; similarly for $\mathbf{p}_1', \mathbf{p}_1''$. Let $P_i = (\mathbf{p}_i', \mathbf{p}_i'')$ be the matrix with column vectors $\mathbf{p}_i', \mathbf{p}_i''$. The point $q_0(\mathbf{t})$ on Q_0 with coordinates $\mathbf{t}$ is then given by

$$q_0(\mathbf{t}) = P_0\mathbf{t} + \tfrac{1}{2}\mathbf{n}_0(K_0\mathbf{t}, \mathbf{t}) + O(|\mathbf{t}|^3), \tag{18}$$

where K_0 is the curvature tensor of Q_0 at† $q_0 \equiv q_0(0)$. For $q_1(\mathbf{s})$ we have a similar expression (here $l = |q_1(0) - q_0(0)|$):

$$q_1(\mathbf{s}) = l\mathbf{e}_3 + P_1\mathbf{s} + \tfrac{1}{2}\mathbf{n}_1(K_1\mathbf{s}, \mathbf{s}) + O(|\mathbf{s}|^3). \tag{19}$$

Letting $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the orthonormal basis of the axes x_i in $\mathbb{R}^3$ we get

$$P_0\mathbf{t} = t_1 \cos\alpha \mathbf{e}_1 + t_2\mathbf{e}_2 + t_1 \sin\alpha \mathbf{e}_3$$
$$P_1\mathbf{s} = s_1\mathbf{e}_1 + s_2 \cos\beta \mathbf{e}_2 + s_2 \sin\beta \mathbf{e}_3,$$

† We abuse notation by letting Q_0 be both a patch in $\mathbb{R}^3$ and a rectangle in the coordinate plane.

where α is the angle between $T_{q_0} Q_0$ and the x_1x_2-plane. Similarly, β is the angle between $T_{q_1} Q_1$ and the x_1x_2-plane. Expanding (18) and (19) in the basis $\{e_i\}$, using the notation $\mathcal{K}_0 = (K_0\mathbf{t}, \mathbf{t})$, $\mathcal{K}_1 = (K_1\mathbf{s}, \mathbf{s})$ and $n_{ij} = (\mathbf{n}_i, \mathbf{e}_j)$, we obtain,

$$\begin{aligned} q_1(\mathbf{s}) - q_0(\mathbf{t}) &= [\ell + s_2 \sin\beta - t_1 \sin\alpha + \tfrac{1}{2} n_{13}\mathcal{K}_1 - \tfrac{1}{2} n_{03}\mathcal{K}_0]\mathbf{e}_3 \\ &\quad + [s_1 - t_1 \cos\alpha + \tfrac{1}{2} n_{11}\mathcal{K}_1 - \tfrac{1}{2} n_{01}\mathcal{K}_0]\mathbf{e}_1 \\ &\quad + [s_2 \cos\beta - t_2 + \tfrac{1}{2} n_{12}\mathcal{K}_1 - \tfrac{1}{2} n_{02}\mathcal{K}_0]\mathbf{e}_2 + O(|t|^3 + |s|^3), \end{aligned}$$

so that

$$L(\mathbf{t}, \mathbf{s}) = |q_2(\mathbf{s}) - q_1(\mathbf{t})| = L_0 + L_1 + L_2 + O(|t|^3 + |s|^3),$$

where L_i are the terms of order i in s and t, and the quadratic term, which interests us, is

$$\begin{aligned} L_2 = D^2 L(\mathbf{t}, \mathbf{s}) &= \frac{1}{2}(n_{13}\mathcal{K}_1 - n_{03}\mathcal{K}_0) \\ &\quad + \frac{1}{2\ell}[(s_1 - t_1\cos\alpha)^2 + (s_2\cos\beta - t_2)^2 + (s_2\sin\beta - t_1\sin\alpha)^2]. \end{aligned}$$

With the above choice of the cylindrical patches Q_i we have $\mathcal{K}_0 = (K_0\mathbf{t}, \mathbf{t}) = (1/\varepsilon)t_1^2$, $\mathcal{K}_1 = (K_1\mathbf{s}, \mathbf{s}) = (1/\varepsilon)s_2^2$. Moreover, there exists a constant $c > 0$ such that for all $q_i \in Q_i$ we have $n_{13} \geq c$, $-n_{03} \geq c$, $|\cos\alpha| \geq \frac{1}{2}$ and $|\cos\beta| \geq \frac{1}{2}$. This shows that the quadratic form L_2 satisfies the lower bound as claimed in the lemma. □

Proof of Lemma D. Consider the 4D cube $Q = Q_0 \times Q_1 = \{(\mathbf{t}, \mathbf{s}) : |t_i| \leq \bar{t}_i,\ |s_i| \leq \bar{s}_i,\ i, j = 1, 2\}$, and pick any $q \in \partial Q$. There are four possibilities: q can lie on a 3D face of Q, on a 2D edge, on a 1D edge or on a vertex, depending on whether one, two, three, or all four bounds

$$|t_i| \leq \bar{t}_i, \quad |s_j| \leq \bar{s}_j \tag{19}$$

are reached. In the first of these cases we define the vector $\mathbf{n} = \mathbf{n}(q)$ to be the outward unit normal to ∂Q. In the remaining cases of q lying on a lower-dimensional edge, this edge is an intersection of two or more 3D faces; to each of these faces there corresponds a unit vector as described above and we simply choose $\mathbf{n}(q)$ as the sum of these vectors. It suffices to prove the bound (4) for $q \in$ 3D-face, so that we assume that precisely one of the bounds (19) is reached.

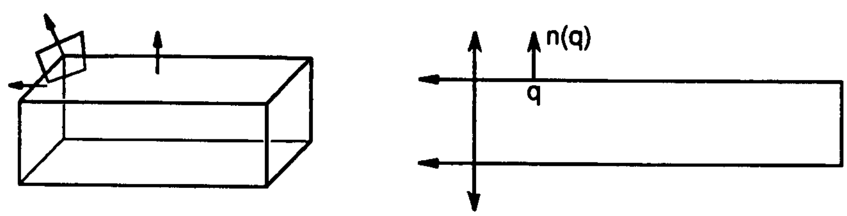

FIGURE 8. Proof of Lemma D.

There are only two essentially distinct possibilities: Case 1 where $t_1 = \bar{t}_1$ and Case 2 where $t_2 = \bar{t}_2$. All the others are trivial modifications of these two, and we now proceed to prove (4) in these two cases.

Case 1. $t_1 = \bar{t}_1$. We have, for $L(q) = |q_1 - q_0|$†

$$D_n L(q) = \frac{\partial}{\partial t_1} L(t_1, t_2, s_1, s_2)_{t_1 = \bar{t}_1}.$$

It is convenient to consider $L^2(q) = (q_1 - q_0)^2$ instead of $L(q)$:

$$\frac{1}{2}\frac{\partial}{\partial t_1} L^2(q) = \frac{1}{2}\frac{\partial}{\partial t_1}(q_1 - q_0)^2 = (q_1 - q_0)\frac{\partial}{\partial t_1}(q_1 - q_0) = (q_0 - q_1)\frac{\partial}{\partial t_1} q_0. \tag{20}$$

Substituting

$$\frac{\partial q_0}{\partial t_1} = \cos\bar{\alpha}\mathbf{e}_2 - \sin\bar{\alpha}\mathbf{e}_3, \quad \text{where } \bar{\alpha} = \varepsilon^{-1}\bar{t}_1 = \frac{\pi}{4}$$

and

$$q_0 - q_1 = -(\ell + O(\varepsilon))\mathbf{e}_3 + O(\sqrt{\varepsilon})\mathbf{e}_1 + O(\sqrt{\varepsilon})\mathbf{e}_2 \tag{21}$$

into (20) we obtain

$$\frac{1}{2}\frac{\partial}{\partial t_1} L^2(q) = \ell \sin\bar{\alpha} + O(\sqrt{\varepsilon}) > \frac{1}{2}\ell,$$

which shows that (4) is easily satisfied in Case 1.

Case 2. $t_2 = \bar{t}_2$. We have

$$\frac{1}{2} D_n L^2(q) = \frac{1}{2}\frac{\partial}{\partial t_2} L^2(q) = (q_0 - q_1)\frac{\partial}{\partial t_2} q_0. \tag{22}$$

Substituting

$$\frac{\partial}{\partial t_2} q_0 = e_2,$$
$$q_0 = \bar{t}_2\mathbf{e}_2 + O(\varepsilon) = \sqrt{\varepsilon}\mathbf{e}_2 + O(\varepsilon),$$
$$q_1 = \ell\mathbf{e}_3 + s_1\mathbf{e}_1 + O(\varepsilon),$$

into (22) we get

$$\tfrac{1}{2} D_n L^2(q) = \sqrt{\varepsilon} + O(\varepsilon) > \tfrac{1}{2}\sqrt{\varepsilon},$$

which proves (4) in Case 2 and thus completes the proof of Lemma D. □

Proof of Lemma A. The proof is by a direct calculation. It is based on the fact that with our choice of the metric the geodesics are governed by linear equations. According to the Maupertuis principle, the geodesics in a conformal metric $\rho(x)$ coincide with the trajectories of a classical particle in a potential field:

$$\ddot{\mathbf{x}} = -\nabla V(\mathbf{x}), \quad V(\mathbf{x}) = E - \rho^2(\mathbf{x}), \tag{23}$$

with the energy

$$\frac{\dot{\mathbf{x}}^2}{2} + V(\mathbf{x}) = E. \tag{24}$$

With our choice of $\rho(x)$,

$$V = E - \rho^2 = E - \frac{1}{\varepsilon^2}(\varepsilon^4 + (1 - \varepsilon^2)r^2), \quad r^2 = x_2^2 + x_3^2, \quad |r| \le \varepsilon,$$

† Here we abuse notation by letting q denote both a point in $\mathbb{R}^3$ and a pair of coordinates on a patch.

and the geodesics inside $T_1(\varepsilon) = \{\mathbf{x} : x_2^2 + x_3^2 \leq \varepsilon^2\}$ are governed by

$$\ddot{x}_1 = 0 \tag{25}$$

$$\ddot{x}_i = 2\frac{1-\varepsilon^2}{\varepsilon^2}x_i, \quad i = 2, 3, \tag{26}$$

together with the energy relation

$$\frac{\dot{\mathbf{x}}^2}{2} = \varepsilon^2 + \frac{1-\varepsilon^2}{\varepsilon^2}r^2, \quad |r| \leq \varepsilon. \tag{27}$$

It suffices to show that the boundary value problem (25)–(27) with boundary conditions

$$\mathbf{x}^{\text{start}} \equiv p = (0, 0, \varepsilon), \quad \mathbf{x}^{\text{end}} \equiv q = (a, \varepsilon\cos\theta, \varepsilon\sin\theta) \tag{28}$$

has a unique solution.

Any solution to (25)–(26) with $\mathbf{x}(0) = \mathbf{x}^{\text{start}}$ is of the form (here v_1, v_2, v_3 are constants):

$$x_1(t) = v_1 t \tag{29x}$$

$$x_2(t) = \delta v_2 \sinh\delta^{-1}t, \quad \delta = \frac{\varepsilon}{\sqrt{2(1-\varepsilon^2)}} \sim \frac{\varepsilon}{\sqrt{2}}, \tag{29y}$$

$$x_3(t) = \varepsilon\cosh\delta^{-1}t + \delta v_3 \sinh\delta^{-1}t \tag{29z}$$

and the energy relation (27) gives

$$\frac{v_1^2}{2} + \frac{v_2^2}{2} + \frac{v_3^2}{2} = 1. \tag{29E}$$

We now show that for any $\mathbf{x}^{\text{end}}$ given by (28) with $a \geq \frac{1}{8}$ there exists a unique choice of parameters (v_1, v_2, v_3, T) subject to (29E) such that $\mathbf{x}(T) = \mathbf{x}^{\text{end}}$. This latter boundary condition is

$$v_1 T = a \tag{30x}$$

$$\delta v_2 \sinh\delta^{-1}T = \varepsilon\cos\theta \tag{30y}$$

$$\varepsilon\cosh\delta^{-1}T + \delta v_3 \sinh\delta^{-1}T = \varepsilon\sin\theta. \tag{30z}$$

Expressing $v_1 = a/T$, we use (30y) and (30z) to find

$$v_2^2 + v_3^2 = \frac{\varepsilon^2}{\delta^2}\left(\sinh^2\frac{T}{\delta}\right)^{-1}\left(1 - 2\sin\theta\cosh\frac{T}{\delta} + \cosh^2\frac{T}{\delta}\right).$$

Substituting these into the energy relation (29E) we get

$$\frac{a^2}{T^2} + \frac{\varepsilon^2}{\delta^2}\frac{\cosh^2(T/\delta) - 2\sin\theta\cosh(T/\delta) + 1}{\sinh^2(T/\delta)} = 2. \tag{31}$$

Since the second term in the left-hand side is positive, we must have $(a^2/T^2) \leq 2$, i.e. $T \geq a/\sqrt{2} = 1/8\sqrt{2}$. For all such T the left-hand side is a monotone decreasing function (decreasing from a value > 2 at $T = a/\sqrt{2}$ to $\varepsilon^2/\delta^2 = 2(1-\varepsilon^2) < 2$ at $T = \infty$). Thus there exists a unique T such that (31) holds. Once T has been found, $v_{1,2,3}$ are defined uniquely by (30). The energy relation (29E) is automatically satisfied. □

Proof of Lemma B. The smallness of DL^{ε} is due to the fact that the geodesics in question are nearly normal to ∂T. To make this more precise, we remove the restriction $q \in \partial T$, obtaining an extension $\mathcal{L}$ of L. To be specific, let r be the distance to ∂T; we have $\mathcal{L} = \mathcal{L}(s_1, s_2, s_3, s_4, r_1, r_2)$ with $\mathcal{L}(s_1, s_2, s_3, s_4, 0, 0) = L(s_1, s_2, s_3, s_4)$. This extension $\mathcal{L}$ is well defined for all $r_1,\ r_2$ small enough and for $|s_1 - s_3| \geq \frac{1}{8}$.

We will prove

$$|\nabla \mathcal{L}(s, r) - n(s)|_{r_1=r_2=0} = O(\varepsilon) \tag{32}$$

and

$$|\partial_{s_i}\big(\nabla \mathcal{L}(s, r) - n(s)\big)|_{r_1=r_2=0} = O(\varepsilon), \quad i = 1, 2. \tag{33}$$

However, we first show that (32) and (33) imply the desired estimates on L.

Letting τ_i be the unit tangent vector to ∂T corresponding to $\partial/\partial s_i$, we have

$$\partial_{s_i} L(s) = \nabla \mathcal{L} \cdot \tau_i = (\nabla \mathcal{L} - n) \cdot \tau_i + n \cdot \tau_i = (\nabla \mathcal{L} - n) \cdot \tau_i = O(\varepsilon),$$

proving $|DL| = O(\varepsilon)$. Next, we obtain

$$\begin{aligned} \partial_{s_i s_j} L(s) &= \partial_{s_i}(\nabla \mathcal{L} \cdot \tau_j) = (\partial_{s_i} \nabla \mathcal{L}) \cdot \tau_j + \nabla \mathcal{L} \cdot \partial_{s_i} \tau_j \\ &= (\partial_{s_i} n) \cdot \partial_{s_i} \tau_j + O(\varepsilon) + n \cdot \partial_{s_i} \tau_j + O(\varepsilon) = \partial_{s_i}(n \cdot \tau_j) + O(\varepsilon) = O(\varepsilon). \end{aligned}$$

where (33) and (32) were used. It remains to prove (32) and (33).

To prove (32) and (33) we note that since ρ is a conformal metric, we have

$$\nabla \mathcal{L}(q) = \frac{1}{|\dot{x}|} \dot{x},$$

in the notation of the proof of Lemma A. From (29)

$$\dot{x}_1(T) = v_1, \quad \dot{x}_2(T) = v_2 \cosh \frac{T}{\delta}, \quad \dot{x}_3 = \frac{\varepsilon}{\delta} \sinh \frac{T}{\delta} + v_3 \cosh \frac{T}{\delta} \tag{34}$$

and from (30) we get

$$\dot{x}_1(T) = \frac{a}{T}, \quad \dot{x}_2(T) = \cos\theta \cdot A, \quad \dot{x}_3(T) = \sin\theta \cdot A, \tag{35}$$

where

$$A = \frac{\varepsilon}{\delta} \coth \frac{T}{\delta}.$$

From (31) we get an estimate on T:

$$T = \frac{a}{\varepsilon\sqrt{2}} + O(e^{-c/\varepsilon}),$$

and using it in (2) we get $\dot{x}_1(T) = O(\varepsilon)$. The energy relation $\dot{x}^2(T) = 2$ applied to (35) gives $A = \sqrt{2} + O(\varepsilon)$; this shows that (recalling that $|\dot{x}(T)| = \sqrt{2}$)

$$\frac{1}{\sqrt{2}} \dot{x}(T) - \mathbf{n}(x(T)) = O(\varepsilon),$$

proving (32).

To prove (33) we note that the time $T(s_1, s_2)$ depends on the coordinates s_1, s_2 mentioned above, as do v_1, v_2, v_3 and n. We have to prove

$$\partial_{s_i}\left(\frac{1}{\sqrt{2}}\dot{x}(T(s_1, s_2)) - \mathbf{n}(s_1, s_2)\right) = O(\varepsilon).$$

This is again a straightforward consequence of the above estimates (34), (29), (30), (31) and we omit the details.

So far we have estimated only the derivatives with respect to the coordinates s_1, s_2 of q, but these imply similar estimates on *all* derivatives up to second order in s_1, s_2, s_3, s_4 (here s_3, s_4 denote the coordinates of p), as follows from the translation invariance of $L(s_1, s_2, s_3, s_4)$. For instance,

$$\partial^2_{s_1 s_4} L(s_1, s_2, s_3, s_4) = \partial^2_{s_1 s_4} L(s_1, s_2 - s_4, s_3, 0) = -\frac{\partial^2}{\partial s_1 \partial s_2'} L(s_1, s_2', s_3, 0) = O(\varepsilon).$$

Acknowledgement. It is a pleaure to thank Jürgen Moser for his encouragement in writing down these ideas and for many stimulating conversations.

REFERENCES

[A] S. Aubry and P. Y. LeDaeron. The discrete Frenkel–Kontorova model and its extensions. *Physica* **8D** (1983), 381–422.

[B1] V. Bangert. Minimal geodesics. *Ergod. Th. & Dynam. Sys.* **10** (1990), 263–286.

[B2] V. Bangert. Mather sets for twist maps and geodesics on tori. *Dynamics Reported*, vol. 1. Eds U. Kirchgraber and H. P. Walther. Wiley and Teubner, Chichester–Stuttgart, 1988, pp. 1–56.

[BK] D. Bernstein and A. Katok. Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians. *Invent. Math.* **88** (1987), 225–241.

[BP] M. Byalyi and L. Polterovich. Geodesic flows on the two-dimensional torus and phase transitions 'commensurability–noncommensurability'. *Functional Anal. Appl.* **20** (1986), 260–266.

[C] C. C. Conley. Isolated invariant sets and the Morse index. CBMS, No 38, AMS, 1978.

[FK] V. Frenkel and T. Kontorova. **ZhETF 8** (1938), 364.

[G] C. Gole. A new proof of the Aubry–Mather theorem. *ETH preprint*, 1991.

[H] G. A. Hedlund. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. Math.* **33** (1932), 719–739.

[K] A. Katok. Some remarks on Birkhoff and Mather twist theorems. *Ergod. Th. & Dynam. Sys.* **2** (1982), 185–194.

[M] R. Mañé. On minimizing measures of Lagrangian dynamical systems. *Nonlinearity* **5** (1992), 623–638.

[M2] R. Mañé. Global variational methods in conservative dynamics. IMPA, Rio de Janeiro, 1993.

[MA1] J. Mather. Existence of quasi-periodic orbits for twist homeomorphisms of the annulus. *Topology* **21** (1982), 457–467.

[MA2] J. Mather. Action-minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.* **207** (1991), 169–207.

[MO1] J. Moser. Monotone twist mappings and the calculus of variations. *Ergod. Th. & Dynam. Sys.* **6** (1986), 401–413.

[MO2] J. Moser. *Break-down of Stability, Nonlinear Dynamics Aspects of Particle Accelerators (Lecture Notes in Physics 247).* Springer, 1986, pp. 492–518.

[MO3] J. Moser. Recent developments in the theory of dynamical systems. *SIAM Review* **28** (1986), 459–485.