Comment. Math. Helv. 90 (2015), 225–254
DOI 10.4171/CMH/352

Commentarii Mathematici Helvetici

Periodic billiard trajectories and Morse theory on loop spaces

Kei Irie

Abstract. We study periodic billiard trajectories on a compact Riemannian manifold with boundary by applying Morse theory to Lagrangian action functionals on the loop space of the manifold. Based on the approximation method proposed by Benci–Giannoni, we prove that nonvanishing of relative homology of a certain pair of loop spaces implies the existence of a periodic billiard trajectory. We also prove a parallel result for path spaces. We apply those results to show the existence of short billiard trajectories and short geodesic loops. Further, we recover two known results on the length of a shortest periodic billiard trajectory in a convex body: Ghomi's inequality, and Brunn–Minkowski type inequality proposed by Artstein-Avidan–Ostrover.

Mathematics Subject Classification (2010). 37J45, 70H12, 52A20.

Keywords. Periodic billiard trajectory, loop space.

1. Introduction and results

In this section, we describe our main results and the structure of this paper.

1.1. Definitions of periodic and brake billiard trajectories. First, let us fix the definition of a periodic billiard trajectory. We also introduce the notion of a brake billiard trajectory, which is a relative version of the periodic trajectory.

Let Q be a Riemannian manifold with C^∞ boundary. We set $S^1 := \mathbb{R}/\mathbb{Z}$. A nonconstant, continuous, and piecewise C^∞ map $\gamma : S^1 \to Q$ is called a *periodic billiard trajectory* if there exists a finite set $\mathcal{B}_\gamma \subset S^1$ such that $\ddot{\gamma} \equiv 0$ on $S^1 \setminus \mathcal{B}_\gamma$, and every $t \in \mathcal{B}_\gamma$ satisfies the following conditions:

B-(i): $\gamma(t) \in \partial Q$.

B-(ii): $\dot{\gamma}^{\pm}(t) := \lim_{h \to 0\pm} \dot{\gamma}(t+h)$ satisfies the following equation:

$$\dot{\gamma}^+(t) + \dot{\gamma}^-(t) \in T_{\gamma(t)}\partial Q, \quad \dot{\gamma}^+(t) - \dot{\gamma}^-(t) \in (T_{\gamma(t)}\partial Q)^\perp \setminus \{0\}.$$

This equation is called the *law of reflection*.

Remark 1.1. Here, we state some remarks on the above definition.

- A periodic billiard trajectory γ might be a closed geodesic on Q. In that case, $\mathcal{B}_\gamma = \emptyset$.
- If γ is tangent to ∂Q at $\gamma(t)$, B-(ii) does not hold since $\dot{\gamma}^+(t) - \dot{\gamma}^-(t) = 0$. Therefore, $\gamma^{-1}(\partial Q)$ might be strictly larger than $\mathcal{B}_\gamma$.
- The law of reflection implies that $|\dot{\gamma}|$ is constant on $S^1 \setminus \mathcal{B}_\gamma$. Moreover, $|\dot{\gamma}| \neq 0$ since γ is a nonconstant map.

A nonconstant, continuous, and piecewise C^∞ map $\gamma : [0,1] \to Q$ is called a *brake billiard trajectory* if it satisfies the following conditions:

- There exists a finite set $\mathcal{B}_\gamma \subset (0,1)$ such that $\ddot{\gamma} \equiv 0$ on $[0,1] \setminus \mathcal{B}_\gamma$, and every $t \in \mathcal{B}_\gamma$ satisfies B-(i), B-(ii).
- $\gamma(0), \gamma(1) \in \partial Q$. $\dot{\gamma}^+(0), \dot{\gamma}^-(1)$ are perpendicular to ∂Q.

The name "brake" billiard trajectory is derived from the notion of a brake orbit in classical mechanics (see [11] pp.131). In both (periodic and brake) cases, elements of $\mathcal{B}_\gamma$ are called *bounce times* of γ.

For any brake billiard trajectory $\gamma : [0,1] \to Q$, we have a periodic billiard trajectory $\Gamma : S^1 \to Q$, which is defined as

$$\Gamma(t) := \begin{cases} \gamma(2t) & (0 \le t \le 1/2) \\ \gamma(2-2t) & (1/2 \le t \le 1). \end{cases}$$

This is a genuine billiard trajectory, i.e., $\mathcal{B}_\Gamma \neq \emptyset$. If γ satisfies $\mathcal{B}_\gamma = \emptyset$, Γ is called a *bouncing ball orbit*.

1.2. Billiard trajectory and topology of path/loop spaces. We state our first result, Theorem 1.2, which claims that the nonvanishing of the relative homology of a certain pair of loop spaces implies the existence of a periodic billiard trajectory. We also prove a parallel result for brake billiard trajectories.

First, we fix some notations. A continuous map $\gamma : S^1 \to Q$ is of class $W^{1,2}$, if it is absolutely continuous and its first derivative is square-integrable. $W^{1,2}(S^1, Q)$ denotes the space of $W^{1,2}$-maps $S^1 \to Q$. $W^{1,2}([0,1], Q)$ is defined in the same manner. We use the following notations:

$$\Lambda(Q) := W^{1,2}([0,1], Q), \quad \Omega(Q) := W^{1,2}(S^1, Q).$$

These spaces are equipped with natural topologies. For any subset $S \subset Q$, we set

$$\Lambda(S) := \{\gamma \in \Lambda(Q) \mid \gamma([0,1]) \subset S\}, \quad \Omega(S) := \{\gamma \in \Omega(Q) \mid \gamma(S^1) \subset S\}.$$

They are equipped with induced topologies as subsets of $\Lambda(Q)$, $\Omega(Q)$.

We define $\mathcal{E} : \Lambda(Q) \to \mathbb{R}$ by $\mathcal{E}(\gamma) := \int_0^1 \frac{|\dot{\gamma}(t)|^2}{2} \, dt$. $\mathcal{E} : \Omega(Q) \to \mathbb{R}$ is defined in the same manner. For any $a \in \mathbb{R}$, we define

$$\Lambda^a(Q) := \{\gamma \in \Lambda(Q) \mid \mathcal{E}(\gamma) < a\}, \quad \Omega^a(Q) := \{\gamma \in \Omega(Q) \mid \mathcal{E}(\gamma) < a\}.$$

When $a < b$, one has obvious inclusions $\Lambda^a(Q) \subset \Lambda^b(Q)$ and $\Omega^a(Q) \subset \Omega^b(Q)$. Let δ be any positive number. We denote the distance on Q by dist, and define

$$\begin{aligned} Q(\delta) &:= \{q \in Q \mid \mathrm{dist}(q, \partial Q) \geq \delta\}, \\ \Lambda_\delta(Q) &:= \Lambda(Q) \setminus \Lambda(Q(\delta)) = \{\gamma \in \Lambda(Q) \mid \mathrm{dist}(\gamma([0,1]), \partial Q) < \delta\}, \\ \Omega_\delta(Q) &:= \Omega(Q) \setminus \Omega(Q(\delta)) = \{\gamma \in \Omega(Q) \mid \mathrm{dist}(\gamma(S^1), \partial Q) < \delta\}. \end{aligned}$$

When $\delta' < \delta$, one has obvious inclusions $\Lambda_{\delta'}(Q) \subset \Lambda_\delta(Q)$ and $\Omega_{\delta'}(Q) \subset \Omega_\delta(Q)$.

Theorem 1.2. *Let Q be a compact Riemannian manifold with C^∞ boundary, $a < b$ be positive real numbers, and j be a non-negative integer.*

(i): *If* $\varprojlim_{\delta \to 0} H_j(\Lambda^b(Q) \cup \Lambda_\delta(Q), \Lambda^a(Q) \cup \Lambda_\delta(Q)) \neq 0$, *there exists a brake billiard trajectory γ on Q such that $\sharp\mathcal{B}_\gamma \leq j-2$ and* $\mathrm{length}(\gamma) \in [\sqrt{2a}, \sqrt{2b}]$.

(ii): *If* $\varprojlim_{\delta \to 0} H_j(\Omega^b(Q) \cup \Omega_\delta(Q), \Omega^a(Q) \cup \Omega_\delta(Q)) \neq 0$, *there exists a periodic billiard trajectory γ on Q such that $\sharp\mathcal{B}_\gamma \leq j$ and* $\mathrm{length}(\gamma) \in [\sqrt{2a}, \sqrt{2b}]$.

Remark 1.3. Let us verify Theorem 1.2 when Q is a closed manifold. In this case, $H_*(\Lambda^b(Q), \Lambda^a(Q)) = 0$ always holds, and therefore, the assumption of (i) is never satisfied. On the other hand, (ii) claims that if $H_*(\Omega^b(Q), \Omega^a(Q)) \neq 0$, then there exists a closed geodesic γ on Q such that $\mathrm{length}(\gamma) \in [\sqrt{2a}, \sqrt{2b}]$. This is a well-known fact in the study of closed geodesics (see e.g., [13]). Thus, the main point of Theorem 1.2 is when Q has a nonempty boundary, and one can think of it as the billiard version of the above-mentioned classical fact.

We explain the idea of the proof of Theorem 1.2. For simplicity, we only discuss case (i). We take a "potential function" $U : \mathrm{int}Q \to \mathbb{R}_{\geq 0}$ which diverges to ∞ near ∂Q. We also take $\varepsilon > 0$, and study the following equation for $\gamma : [0,1] \to \mathrm{int}Q$.

$$\dot{\gamma}(0) = \dot{\gamma}(1) = 0, \quad \ddot{\gamma}(t) + \varepsilon \nabla U(\gamma(t)) \equiv 0. \tag{1.1}$$

As is well-known, the solutions of this equation are critical points of the Lagrangian functional $\mathcal{L}_\varepsilon$ on the path space $\Lambda(\mathrm{int}Q)$, which is defined as

$$\mathcal{L}_\varepsilon(\gamma) := \int_0^1 \frac{|\dot{\gamma}(t)|^2}{2} - \varepsilon U(\gamma(t)) \, dt.$$

Proposition 2.2, which is proved in Section 2, shows that one can prove the existence of a solution of (1.1) using Morse theory for the functional $\mathcal{L}_\varepsilon$.

Suppose that we have a solution γ_ε of (1.1) for any sufficiently small $\varepsilon > 0$, which satisfies certain estimates on $\mathcal{L}_\varepsilon(\gamma_\varepsilon)$ and the Morse index. Proposition 3.1, which is proved in Section 3, claims that we can get a billiard trajectory γ as a limit of γ_ε as $\varepsilon \to 0$, which satisfies corresponding estimates on $\mathrm{length}(\gamma)$ and $\sharp\mathcal{B}_\gamma$. Combining the results in Sections 2 and 3, we will complete the proof of Theorem 1.2 in Section 4.

The above strategy of the proof is significantly influenced by [7]. In particular, our arguments in Sections 2 and 3 closely follow the arguments in [7]. Nevertheless, we explain most details for the reader's convenience.

1.3. Short billiard trajectory. As an application of Theorem 1.2, we prove the existence of short billiard trajectories. First, we state the result. Let Q be a compact, connected Riemannian manifold with a nonempty C^∞ boundary. $r(Q)$ denotes the *inradius* of Q, i.e., $r(Q) := \max_{q\in Q} \mathrm{dist}(q, \partial Q)$. It is easy to see that $r(Q) < \infty$.

Theorem 1.4. *Let j be a positive integer such that $H_j(Q, \partial Q : \mathbb{Z}) \neq 0$. Then, there exist the following billiard trajectories on Q:*

- *A brake billiard trajectory γ_B, such that $\sharp\mathcal{B}_{\gamma_B} \le j-1$ and $\mathrm{length}(\gamma_B) \le 2jr(Q)$.*
- *A periodic billiard trajectory γ_P, such that $\sharp\mathcal{B}_{\gamma_P} \le j+1$ and $\mathrm{length}(\gamma_P) \le 2(j+1)r(Q)$.*

Remark 1.5. To the best of the author's knowledge, there are very few examples in which the above estimates are sharp. It is easy to see that the estimates are sharp for $j = 1$, considering the case Q is a line segment. For $j = 2$, the estimates $\sharp\mathcal{B}_{\gamma_B} \le 1$ and $\sharp\mathcal{B}_{\gamma_P} \le 3$ are sharp, since there exists a planar domain that does not contain any bouncing ball orbits, see Figure 1-(b) in [10].

Theorem 1.4 is proved in Section 5. In this subsection, we explain some consequences of Theorem 1.4. Let us introduce the following notations.

$$\mu_B(Q) := \inf\{\mathrm{length}(\gamma) \mid \gamma : \text{brake billiard trajectory on } Q\},$$
$$\mu_P(Q) := \inf\{\mathrm{length}(\gamma) \mid \gamma : \text{periodic billiard trajectory on } Q\}.$$

As an immediate consequence of Theorem 1.4, we obtain the following estimate.

Corollary 1.6. *Let n denote the dimension of Q. Then, there holds $\mu_B(Q) \le 2nr(Q)$ and $\mu_P(Q) \le 2(n+1)r(Q)$.*

The above estimate of μ_P was already proved in [5] for convex domains in $\mathbb{R}^n$, and in [12] for arbitrary domains in $\mathbb{R}^n$. For other previous results on short periodic billiard trajectories, see [5] Section 1, and the references therein.

Another consequence of Theorem 1.4 is a new proof of the following result on short geodesic loops, which is proved in [14]. The original proof in [14] is based on the Birkhoff curve shortening process, and it seems considerably different from our arguments.

Corollary 1.7 (Rotman [14]). *Let M be a closed Riemannian manifold, $p \in M$, and j be a positive integer. If $\pi_j(M, p) \neq 0$, there exists a nonconstant geodesic loop γ at p (i.e., a geodesic path $\gamma : [0,1] \to M$ such that $\gamma(0), \gamma(1) = p$) such that* $\text{length}(\gamma) \leq 2j\,\text{diam}(M)$.

We prove Corollary 1.7 by considering a short brake billiard trajectory on $\{x \in M \mid \text{dist}(x, p) \geq \varepsilon\}$ and letting $\varepsilon \to 0$. The details will be explained in Section 5.3.

1.4. Length of the shortest periodic billiard trajectory in a convex body. A convex set $K \subset \mathbb{R}^n$ is called a *convex body* if K is compact and $\text{int}K \neq \emptyset$. It is possible to show that, for any convex body K with C^∞ boundary, there exists a periodic billiard trajectory in K of length $\mu_P(K)$ (see Remark 6.5).

Let us recall two remarkable geometric inequalities on μ_P of convex bodies, which are proved in [5] and [10]. In Section 6, we recover these results as applications of our method. A recent paper [3] obtained similar proofs based on the results in [8] in a more general setting of Finsler billiards.

The first inequality is the Brunn–Minkowski type inequality [5]. For any two convex bodies $K_1, K_2 \subset \mathbb{R}^n$, their *Minkowski sum* $K_1 + K_2 := \{x_1 + x_2 \mid x_1 \in K_1, x_2 \in K_2\}$ is a convex body. The following result is proved in [5], based on their Brunn–Minkowski type inequality for symplectic capacity [4].

Theorem 1.8 (Artstein-Avidan–Ostrover [5]). *Let K_1, K_2 be convex bodies in $\mathbb{R}^n$. Suppose that K_1, K_2, and $K_1 + K_2$ have C^∞ boundaries. Then,*

$$\mu_P(K_1 + K_2) \geq \mu_P(K_1) + \mu_P(K_2).$$

Equality holds if and only if there exists a closed curve which, up to parallel displacement and scaling, is the shortest periodic billiard trajectory of both K_1 and K_2.

The second inequality is a lower bound of μ_P by inradius, which is proved in [10] by beautiful elementary arguments. $\text{width}(K)$ denotes the thickness of the narrowest slab that contains K.

Theorem 1.9 (Ghomi [10]). *For any convex body $K \subset \mathbb{R}^n$ with C^∞ boundary, there holds $\mu_P(K) \geq 4r(K)$. Equality holds if and only if $2r(K) = \text{width}(K)$. In this case, every shortest periodic billiard trajectory in K is a bouncing ball orbit.*

Remark 1.10. We only partially recover the original results in [5] and [10]. In [5], the authors prove Theorem 1.8 in a more general setting of Minkowski billiards, whereas we discuss only Euclidean billiards. On the other hand, [10] does not assume the smoothness of ∂K.

2. Approximating problem

In this section, we study an approximating problem for the billiard problem, which was introduced in [7]. In Section 2.1, we fix the setting and state Proposition 2.2, which is the main result in this section. Section 2.2 is devoted to its proof.

Throughout Sections 2 and 3, Q denotes a compact, connected Riemannian manifold with a nonempty boundary. We abbreviate $\Lambda(\mathrm{int}Q)$, $\Omega(\mathrm{int}Q)$ as Λ, Ω. These spaces have natural structures of smooth Hilbert manifolds. For any $\gamma_\Lambda \in \Lambda$ and $\gamma_\Omega \in \Omega$, tangent spaces at γ_Λ and γ_Ω are described as

$$T_{\gamma_\Lambda}\Lambda = W^{1,2}([0,1], \gamma_\Lambda^*(TQ)), \quad T_{\gamma_\Omega}\Omega = W^{1,2}(S^1, \gamma_\Omega^*(TQ)).$$

2.1. Setting. We take and fix $\rho \in C^\infty(\mathbb{R}_{\geq 0})$ such that

- $\rho(t) = t$ for any $0 \leq t \leq 1$.
- $0 \leq \rho(t) \leq 2$, $0 \leq \rho'(t) \leq 1$ for any $t \geq 0$.
- $\rho(t) = 2$ for any $t \geq 3$.

We define $d \in C^0(Q)$ by $d(q) := \mathrm{dist}(q, \partial Q)$. Recall the notation $Q(\delta) := \{q \in Q \mid d(q) \geq \delta\}$ in Section 1.2. When $\delta > 0$ is sufficiently small, d is of C^∞ and satisfies $|\nabla d| \equiv 1$ on $Q \setminus Q(3\delta)$. For such δ, we define $h_\delta \in C^\infty(Q)$ and $U_\delta \in C^\infty(\mathrm{int}Q)$ by

$$h_\delta(q) := \delta\rho(d(q)/\delta), \quad U_\delta(q) := h_\delta(q)^{-2} - (2\delta)^{-2}.$$

Notice that $U_\delta \equiv 0$ on $Q(3\delta)$. In Sections 2 and 3, we fix δ and abbreviate h_δ, U_δ as h, U. The following lemma is easy to prove, and we will use it a few times.

Lemma 2.1.

(i): *Let ν be a smooth vector field on Q such that $\nu \equiv -\nabla d$ on $Q \setminus Q(3\delta)$. Then,* $|\nabla U(q)| = \langle \nabla U(q), \nu(q)\rangle$ *for any* $q \in \mathrm{int}Q$.

(ii): *There holds* $\lim_{q\to\partial Q} U(q)/|\nabla U(q)| = 0$.

First, we consider the approximating problem for brake billiard trajectories. Suppose that $V \in C^\infty([0,1] \times \mathrm{int}Q)$ satisfies the following property.

V-(i): There exists $\varepsilon > 0$ and a compact set $K \subset \mathrm{int}Q$ such that $V(t,q) = \varepsilon U(q)$ for any $t \in [0,1]$, $q \notin K$.

We define $L_V^\Lambda \in C^\infty([0,1] \times T(\mathrm{int}Q))$ and $\mathcal{L}_V^\Lambda : \Lambda \to \mathbb{R}$ by

$$L_V^\Lambda(t,q,v) := \frac{|v|^2}{2} - V(t,q)\ (q \in \mathrm{int}Q, v \in T_qQ), \quad \mathcal{L}_V^\Lambda(\gamma) := \int_0^1 L_V^\Lambda(t,\gamma,\dot\gamma)\,dt.$$

$\mathcal{L}_V^\Lambda$ is a C^∞ functional on Λ, and its differential is computed as

$$d\mathcal{L}_V^\Lambda(\gamma)(\zeta) = \int_0^1 \langle \dot\gamma, \nabla_t\zeta\rangle - dV_t(\zeta)\,dt \quad (\zeta \in T_\gamma\Lambda),$$

where ∇_t denotes the Levi-Civita covariant derivative, and $V_t \in C^\infty(\mathrm{int}Q)$ is defined by $V_t(q) := V(t,q)$. $\gamma \in \Lambda$ satisfies $d\mathcal{L}_V^\Lambda(\gamma) = 0$ if and only if it is of class C^∞ and satisfies

$$\dot\gamma(0) = \dot\gamma(1) = 0, \quad \ddot\gamma(t) + \nabla V_t(\gamma(t)) \equiv 0. \tag{2.1}$$

For any γ satisfying (2.1), the Hessian of $\mathcal{L}_V^\Lambda$ at γ is given by the following formula, where R denotes the curvature tensor.

$$d^2\mathcal{L}_V^\Lambda(\gamma)(\eta,\zeta) = \int_0^1 \langle\nabla_t\eta, \nabla_t\zeta\rangle - \langle R(\dot\gamma,\eta)(\zeta),\dot\gamma\rangle - \langle\nabla_\eta\nabla V_t(\gamma),\zeta\rangle\, dt \quad (\eta,\zeta\in T_\gamma\Lambda). \tag{2.2}$$

$\mathrm{ind}(\gamma)$ denotes the Morse index of γ, which is the number of negative eigenvalues of $d^2\mathcal{L}_V^\Lambda(\gamma)$. As is well known, $\mathrm{ind}(\gamma) < \infty$ (see e.g., [2] Proposition 3.1 (iii)). γ is called nondegenerate if 0 is not an eigenvalue of $d^2\mathcal{L}_V^\Lambda(\gamma)$.

Next, we consider the approximating problem for periodic billiard trajectories. Suppose that $V \in C^\infty(S^1 \times \mathrm{int}Q)$ satisfies the following property.

V-(ii): There exists $\varepsilon > 0$ and a compact set $K \subset \mathrm{int}Q$ such that $V(t,q) = \varepsilon U(q)$ for any $t \in S^1, q \notin K$.

We define $L_V^\Omega \in C^\infty(S^1 \times T(\mathrm{int}Q))$ and $\mathcal{L}_V^\Omega : \Omega \to \mathbb{R}$ by

$$L_V^\Omega(t,q,v) = \frac{|v|^2}{2} - V(t,q), \quad \mathcal{L}_V^\Omega(\gamma) := \int_{S^1} L_V^\Omega(t,\gamma,\dot\gamma)\, dt.$$

$\gamma \in \Omega$ satisfies $d\mathcal{L}_V^\Omega(\gamma) = 0$ if and only if it is of class C^∞ and satisfies $\ddot\gamma(t) + \nabla V_t(\gamma(t)) \equiv 0$. The goal of this section is to prove the following proposition.

Proposition 2.2. *Let $a < b$ be real numbers, and j be a non-negative integer.*

(i): *For any $V \in C^\infty([0,1] \times \mathrm{int}Q)$ that satisfies V-(i) and*

$$H_j(\{\mathcal{L}_V^\Lambda < b\}, \{\mathcal{L}_V^\Lambda < a\}) \neq 0,$$

there exists $\gamma \in \Lambda$ such that $d\mathcal{L}_V^\Lambda(\gamma) = 0$, $\mathcal{L}_V^\Lambda(\gamma) \in [a,b]$, and $\mathrm{ind}(\gamma) \le j$.

(ii): *For any $V \in C^\infty(S^1 \times \mathrm{int}Q)$ that satisfies V-(ii) and*

$$H_j(\{\mathcal{L}_V^\Omega < b\}, \{\mathcal{L}_V^\Omega < a\}) \neq 0,$$

there exists $\gamma \in \Omega$ such that $d\mathcal{L}_V^\Omega(\gamma) = 0$, $\mathcal{L}_V^\Omega(\gamma) \in [a,b]$, and $\mathrm{ind}(\gamma) \le j$.

2.2. Proof of Proposition 2.2. We only prove (i), since (ii) can be proved by parallel arguments. In this subsection, we abbreviate L_V^Λ and $\mathcal{L}_V^\Lambda$ as L_V and $\mathcal{L}_V$, respectively. As a first step, we need the following result.

Lemma 2.3. *Let $(\gamma_j)_j$ be a sequence in Λ, such that* $\lim_{j\to\infty} \mathrm{dist}(\gamma_j([0,1]), \partial Q) = 0$ *and* $\sup_j \|\dot{\gamma}_j\|_{L^2} < \infty$. *Then, there holds* $\lim_{j\to\infty} \int_0^1 h(\gamma_j)^{-2}\, dt = \infty$.

Proof. See Lemma 3.6 in [7]. □

For any $\gamma \in \Lambda$ and $\eta, \zeta \in T_\gamma \Lambda$, we define a Riemannian metric $\langle \cdot, \cdot \rangle_\Lambda$ on Λ as

$$\langle \eta, \zeta \rangle_\Lambda := \int_0^1 \langle \eta, \zeta \rangle + \langle \nabla_t \eta, \nabla_t \zeta \rangle \, dt, \quad \|\eta\|_\Lambda := \langle \eta, \eta \rangle_\Lambda^{1/2}. \tag{2.3}$$

$\|\cdot\|_\Lambda$ defines a distance function d_Λ on Λ in the obvious manner. This metric structure on $\Lambda = \Lambda(\mathrm{int} Q)$ naturally extends to $\Lambda(Q)$, and this makes $\Lambda(Q)$ a complete metric space. (Notice that $\Lambda(Q)$ is *not* a Hilbert manifold, even with boundary.)

Lemma 2.4. *For any interval $J \subset \mathbb{R}$, which is closed and bounded from below, $(\mathcal{L}_V^{-1}(J), d_\Lambda)$ is a complete metric space.*

Proof. Let $(\gamma_j)_j$ be a Cauchy sequence on $(\mathcal{L}_V^{-1}(J), d_\Lambda)$. There exists $\gamma_\infty \in \Lambda(Q)$ such that $\lim_{j\to\infty} d_\Lambda(\gamma_j, \gamma_\infty) = 0$. It is sufficient to show that $\gamma_\infty([0,1]) \subset \mathrm{int} Q$.

Suppose that $\gamma_\infty([0,1])$ intersects ∂Q. Then, $\lim_{j\to\infty} \mathrm{dist}(\gamma_j([0,1]), \partial Q) = 0$. On the other hand, $\sup_j \|\dot{\gamma}_j\|_{L^2} < \infty$, since the convergence in d_Λ implies a convergence in $W^{1,2}$-topology. Hence Lemma 2.3 implies $\lim_{j\to\infty} \mathcal{L}_V(\gamma_j) = -\infty$, contradicting the assumption that $\mathcal{L}_V(\gamma_j) \in J$ for all j. □

Next, we discuss the Palais–Smale (PS) condition for $\mathcal{L}_V$. For each $\gamma \in \Lambda$, we define $\nabla \mathcal{L}_V(\gamma) \in T_\gamma \Lambda$, the *gradient vector* of $\mathcal{L}_V$ at γ, as

$$\langle \nabla \mathcal{L}_V(\gamma), \eta \rangle_\Lambda := d\mathcal{L}_V(\eta) \quad (\forall \eta \in T_\gamma \Lambda).$$

Definition 2.5. Let X be a (possibly infinite-dimensional) Riemannian manifold and $f : X \to \mathbb{R}$ be a smooth function. A sequence $(x_j)_{j=1,2,\ldots}$ is a *PS-sequence* of f, if $(f(x_j))_j$ is bounded and $\lim_{j\to\infty} \|\nabla f(x_j)\| = 0$. f is said to satisfy the *PS-condition*, if any PS-sequence of f contains a convergent subsequence.

We will show that $\mathcal{L}_V$ satisfies the PS-condition. Our argument is based on the following result.

Lemma 2.6. *Let M be a closed Riemannian manifold, and suppose that $\Lambda(M)$ is equipped with a Riemannian metric in the same manner as (2.3). For any $W \in C^\infty([0,1] \times M)$,*

$$\mathcal{L}_W : \Lambda(M) \to \mathbb{R}; \quad \gamma \mapsto \int_0^1 \frac{|\dot{\gamma}(t)|^2}{2} - W(t, \gamma(t)) \, dt$$

satisfies the PS-condition.

Proof. The claim follows from Proposition 3.3 in [2] (see also [6]). □

Since our base manifold $\mathrm{int}Q$ is open and V diverges to ∞ near the boundary, we cannot apply Lemma 2.6 directly. Thus, we need the following lemma.

Lemma 2.7. *If $(\gamma_j)_j$ is a PS-sequence of $\mathcal{L}_V$, then* $\inf_j \mathrm{dist}(\gamma_j([0,1]), \partial Q) > 0$.

Proof. Since V satisfies V-(i), there exists $\varepsilon > 0$ such that $V(t,q) - \varepsilon U(q)$ is compactly supported. Then, $\mathcal{L}_{\varepsilon U}(\gamma_j)$, $\|\nabla \mathcal{L}_{\varepsilon U}(\gamma_j)\|_\Lambda$ are both bounded on j, since $(\gamma_j)_j$ is a PS-sequence of $\mathcal{L}_V$. Let us take ν as in Lemma 2.1 (i). Then, there holds

$$\begin{aligned}\int_0^1 |\varepsilon \nabla U(\gamma_j)|\, dt &= \int_0^1 \langle \varepsilon \nabla U(\gamma_j), \nu(\gamma_j)\rangle\, dt \\ &= -\langle \nabla \mathcal{L}_{\varepsilon U}(\gamma_j), \nu(\gamma_j)\rangle_\Lambda + \int_0^1 \langle \dot\gamma_j, \nabla_t(\nu(\gamma_j))\rangle\, dt.\end{aligned}$$

We can bound RHS using the following obvious inequalities.

$$\|\nu(\gamma_j)\|_{L^2} \le \max_{q\in Q} |\nu(q)|, \qquad \|\nabla_t(\nu(\gamma_j))\|_{L^2} \le \max_{q\in Q} |\nabla\nu(q)| \cdot \|\dot\gamma_j\|_{L^2}.$$

Thus, there exists a constant $M_0 > 0$, which is independent on j, such that

$$\int_0^1 |\varepsilon \nabla U(\gamma_j)|\, dt \le M_0(1 + \|\dot\gamma_j\|_{L^2}^2).$$

By Lemma 2.1 (ii), there exists $M_1 > 0$ such that $U(q) \le |\nabla U(q)|/4M_0 + M_1$ for any $q \in \mathrm{int}Q$. Thus,

$$\begin{aligned}\int_0^1 \varepsilon U(\gamma_j)\, dt &\le \frac{1}{4M_0}\int_0^1 |\varepsilon \nabla U(\gamma_j)|\, dt + \varepsilon M_1 \le \frac{1 + \|\dot\gamma_j\|_{L^2}^2}{4} + \varepsilon M_1 \\ &= \frac{1}{2}\left(\mathcal{L}_{\varepsilon U}(\gamma_j) + \int_0^1 \varepsilon U(\gamma_j)\, dt\right) + \frac{1}{4} + \varepsilon M_1.\end{aligned}$$

Therefore, we obtain

$$\frac{1}{2}\int_0^1 \varepsilon U(\gamma_j)\, dt \le \frac{1}{2}\mathcal{L}_{\varepsilon U}(\gamma_j) + \frac{1}{4} + \varepsilon M_1.$$

Since $\mathcal{L}_{\varepsilon U}(\gamma_j)$ is bounded on j, we obtain $\sup_j \int_0^1 \varepsilon U(\gamma_j)\, dt < \infty$ and $\sup_j \|\dot\gamma_j\|_{L^2} < \infty$. Since $U(q) = h(q)^{-2} - (2\delta)^{-2}$, Lemma 2.3 implies $\inf_j \mathrm{dist}(\gamma_j([0,1]), \partial Q) > 0$. □

Lemma 2.8. *$\mathcal{L}_V$ satisfies the PS-condition.*

Proof. Let $(\gamma_j)_j$ be a PS-sequence of $\mathcal{L}_V$. By Lemma 2.7, there exists a compact submanifold $Q' \subset \mathrm{int}Q$, such that $\gamma_j([0,1]) \subset Q'$ for all j.

It is easy to show that there exists a closed Riemannian manifold M, an isometric embedding $e : Q' \to M$, and $W \in C^\infty([0,1] \times M)$, such that $V(t,q) = W(t, e(q))$

for all $t \in [0,1]$, $q \in Q'$. Then, $\{e(\gamma_j)\}_j$ is a PS-sequence of $\mathcal{L}_W$. Hence, by Lemma 2.6, it has a convergent subsequence. Thus, $(\gamma_j)_j$ also has a convergent subsequence. □

We define the *spectrum* of $\mathcal{L}_V$ as

$$\mathrm{Spec}(\mathcal{L}_V) := \{\mathcal{L}_V(\gamma) \mid d\mathcal{L}_V(\gamma) = 0\} \subset \mathbb{R}.$$

Lemma 2.9. $\mathrm{Spec}(\mathcal{L}_V)$ *is closed in* $\mathbb{R}$ *and has a zero measure.*

Proof. Closedness is immediate since $\mathcal{L}_V$ satisfies the PS-condition. To show that $\mathrm{Spec}(\mathcal{L}_V)$ has a zero measure, we modify the arguments in [15] pp.436.

For each $x \in \mathrm{int}Q$, we define $\gamma_x : [0,1] \to \mathrm{int}Q$ by $\gamma_x(0) = x$, $\dot{\gamma}_x(0) = 0$, and $\ddot{\gamma}_x(t) + \nabla V_t(\gamma_x(t)) \equiv 0$. Then, $f(x) := \mathcal{L}_V(\gamma_x)$ is a C^∞ function on $\mathrm{int}Q$, and $\mathrm{Spec}(\mathcal{L}_V)$ is contained in the set of critical values of f. Hence, our claim follows from the Sard theorem for finite-dimensional manifolds. □

Remark 2.10. The above proof is based on the fact that any critical point γ of $\mathcal{L}_V$ is determined by $\gamma(0) \in \mathrm{int}Q$. This argument does not apply directly to the periodic case, since any solution γ of the Euler–Lagrange equation with the periodic boundary condition is determined by $\gamma(0)$ and $\dot{\gamma}(0)$. To prove that $\mathrm{Spec}(\mathcal{L}_V^\Omega)$ has a zero measure for any $V \in C^\infty(S^1 \times TQ)$, we may apply Lemma 3.8 in [15] directly to a Hamiltonian $H \in C^\infty(S^1 \times T^*Q)$, which is the Legendre transform of V.

The following lemma is a key step in the proof of Proposition 2.2.

Lemma 2.11. *Let* $c_- < c_+$ *be real numbers such that* $c_\pm \notin \mathrm{Spec}(\mathcal{L}_V)$*. We set*

$$\mathcal{C}_{(c_-,c_+)} := \{\gamma \in \Lambda \mid \mathcal{L}_V(\gamma) \in (c_-,c_+),\ d\mathcal{L}_V(\gamma) = 0\}.$$

If $H_j(\{\mathcal{L}_V < c_+\}, \{\mathcal{L}_V < c_-\}) \neq 0$ *and all elements of* $\mathcal{C}_{(c_-,c_+)}$ *are nondegenerate critical points of* $\mathcal{L}_V$*, there exists* $\gamma \in \mathcal{C}_{(c_-,c_+)}$ *such that* $\mathrm{ind}(\gamma) = j$.

Proof. We use the theory developed in [1], Section 2. Let us set

$$\hat{M} := \{\mathcal{L}_V < c_+\}, \quad M := \{c_- < \mathcal{L}_V < c_+\}, \quad f := \mathcal{L}_V|_{\hat{M}}.$$

We take a smooth vector field $\hat{X}$ on $\hat{M}$, which is a negative scalar multiple of ∇f and satisfies the following properties:

$$\|\nabla f(p)\| < 1 \implies \hat{X}(p) = -\nabla f(p), \quad \|\nabla f(p)\| \ge 1 \implies 1 \le \|\hat{X}(p)\| \le 2.$$

Let us examine whether $\hat{M}$, M, f, and $\hat{X}$ satisfy conditions (A1)–(A7) in [1], pp.22–23. (A1) follows from Lemma 2.4, (A6) follows from Lemma 2.8, and (A2)–(A5) are immediate. Since $\hat{X}$ is smooth, (A7) is also achieved by a small perturbation of $\hat{X}$, without violating (A1)–(A6) (See Remark 2.1 [2]). Now our claim follows from Theorem 2.8 in [1]. □

V is said to be *regular* if all critical points of $\mathcal{L}_V$ are nondegenerate. The next Lemma 2.12 shows that regularness can be achieved by compactly supported small perturbations.

Lemma 2.12. *For any V satisfying V-(i), there exists a sequence $(V_m)_{m=1,2,\ldots}$ such that all V_m are regular and satisfy V-(i), and $(V_m)_m$ converges to V in C^∞-topology, i.e., for any $k \geq 0$, $\lim_{m\to\infty} \|V - V_m\|_{C^k([0,1]\times Q)} = 0$ holds (notice that $V - V_m$ extends to a C^∞-function on $[0,1] \times Q$).*

Lemma 2.12 can be proved in a similar manner as Theorem 1.1 in [16]. The setting in [16] is slightly different from ours: in [16], the base Riemannian manifold is closed, and the Lagrangian is parametrized by S^1. However, these differences do not affect the proof. Now, we can finish the proof of Proposition 2.2 (i). As explained at the beginning of this subsection, the proof of (ii) is parallel and omitted.

Proof of Proposition 2.2 (i). First, we consider the case when $a, b \notin \mathrm{Spec}(\mathcal{L}_V)$. Since $\mathrm{Spec}(\mathcal{L}_V)$ is closed, there exists $c > 0$ such that $[a-c, a+c]$ and $[b-c, b+c]$ are disjoint from $\mathrm{Spec}(\mathcal{L}_V)$. By Lemma 2.11, inclusions

$$\{\mathcal{L}_V < a - c\} \subset \{\mathcal{L}_V < a\} \subset \{\mathcal{L}_V < a + c\},$$
$$\{\mathcal{L}_V < b - c\} \subset \{\mathcal{L}_V < b\} \subset \{\mathcal{L}_V < b + c\}$$

induce isomorphisms on homologies. In particular, the homomorphism

$$H_j(\{\mathcal{L}_V < b - c\}, \{\mathcal{L}_V < a - c\}) \to H_j(\{\mathcal{L}_V < b + c\}, \{\mathcal{L}_V < a + c\})$$

induced by inclusion is an isomorphism, and the homologies on both sides are isomorphic to $H_j(\{\mathcal{L}_V < b\}, \{\mathcal{L}_V < a\})$, which is nonzero by our assumption.

Take a sequence $(V_m)_m$ as in Lemma 2.12. For sufficiently large m, we have

$$\{\mathcal{L}_V < a - c\} \subset \{\mathcal{L}_{V_m} < a\} \subset \{\mathcal{L}_V < a + c\},$$
$$\{\mathcal{L}_V < b - c\} \subset \{\mathcal{L}_{V_m} < b\} \subset \{\mathcal{L}_V < b + c\}.$$

Hence, there holds $H_j(\{\mathcal{L}_{V_m} < b\}, \{\mathcal{L}_{V_m} < a\}) \neq 0$. By Lemma 2.11, there exists $\gamma_m \in \Lambda$ such that $d\mathcal{L}_{V_m}(\gamma_m) = 0$, $\mathcal{L}_{V_m}(\gamma_m) \in [a, b]$, and $\mathrm{ind}(\gamma_m) = j$.

Since $\lim_{m\to\infty} \|V_m - V\|_{C^1} = 0$, $(\gamma_m)_m$ is a PS-sequence of $\mathcal{L}_V$. Hence, $(\gamma_m)_m$ has a convergent subsequence, and its limit γ satisfies $d\mathcal{L}_V(\gamma) = 0$ and $\mathcal{L}_V(\gamma) \in [a, b]$. $\mathrm{ind}(\gamma) \leq j$ follows from $\mathrm{ind}(\gamma) \leq \liminf_{m\to\infty} \mathrm{ind}(\gamma_m)$, which easily follows from (2.2).

Finally, we consider the general case, i.e., a and b may be in $\mathrm{Spec}(\mathcal{L}_V)$. Since $\mathrm{Spec}(\mathcal{L}_V)$ has a zero measure, there exist increasing sequences $(a_m)_m$, $(b_m)_m$ such that $a_m, b_m \notin \mathrm{Spec}(\mathcal{L}_V)$ for every m, and $a = \lim_m a_m$, $b = \lim_m b_m$. Then, for sufficiently large m, $H_j(\{\mathcal{L}_V < b_m\}, \{\mathcal{L}_V < a_m\}) \neq 0$. Therefore, there exists γ_m such that $d\mathcal{L}_V(\gamma_m) = 0$, $\mathcal{L}_V(\gamma_m) \in [a_m, b_m]$, and $\mathrm{ind}(\gamma_m) \leq j$. $(\gamma_m)_m$ is a PS-sequence of $\mathcal{L}_V$, and therefore, it has a convergent subsequence. Then, its limit γ satisfies $d\mathcal{L}_V(\gamma) = 0$, $\mathcal{L}_V(\gamma) \in [a, b]$, and $\mathrm{ind}(\gamma) \leq j$. $\square$

Remark 2.13. From the above proof, it is easy to see that γ satisfies $\operatorname{ind}(\gamma) + \operatorname{null}(\gamma) \geq j$, where $\operatorname{null}(\gamma)$ denotes the dimension of the kernel of $d^2\mathcal{L}_V(\gamma)$. The referee suggested that this inequality can be used to provide a lower bound of the number of bounce times of billiard trajectories.

3. Billiard trajectory as a limit

As in the previous section, we fix $\delta > 0$ and use abbreviations $h := h_\delta$, $U := U_\delta$. The goal of this section is to prove Proposition 3.1, which enables us to get a billiard trajectory as a limit of solutions of the approximating problem.

Proposition 3.1. *Let $a < b$ be positive real numbers, and j be a non-negative integer.*

- (i): *Suppose that for any sufficiently small $\varepsilon > 0$, there exists $\gamma_\varepsilon \in \Lambda$ such that $d\mathcal{L}^\Lambda_{\varepsilon U}(\gamma_\varepsilon) = 0$, $\mathcal{L}^\Lambda_{\varepsilon U}(\gamma_\varepsilon) \in [a, b]$, and $\operatorname{ind}(\gamma_\varepsilon) \leq j$. Then, there exists a brake billiard trajectory γ such that $\sharp\mathcal{B}_\gamma \leq j - 2$ and $\operatorname{length}(\gamma) \in [\sqrt{2a}, \sqrt{2b}]$.*
- (ii): *Suppose that for any sufficiently small $\varepsilon > 0$, there exists $\gamma_\varepsilon \in \Omega$ such that $d\mathcal{L}^\Omega_{\varepsilon U}(\gamma_\varepsilon) = 0$, $\mathcal{L}^\Omega_{\varepsilon U}(\gamma_\varepsilon) \in [a, b]$, and $\operatorname{ind}(\gamma_\varepsilon) \leq j$. Then, there exists a periodic billiard trajectory γ such that $\sharp\mathcal{B}_\gamma \leq j$ and $\operatorname{length}(\gamma) \in [\sqrt{2a}, \sqrt{2b}]$.*

We only prove (i), since (ii) can be proved by parallel arguments. In the following arguments, we fix γ_ε for each ε, and abbreviate $\mathcal{L}^\Lambda_{\varepsilon U}$ as $\mathcal{L}_\varepsilon$.

Lemma 3.2. $\lim_{\varepsilon \to 0} \int_0^1 \varepsilon U(\gamma_\varepsilon)\, dt = 0$.

Proof. Let us take ν as in Lemma 2.1 (i). By $\ddot{\gamma}_\varepsilon + \varepsilon \nabla U(\gamma_\varepsilon) \equiv 0$ and $\dot{\gamma}_\varepsilon(0) = \dot{\gamma}_\varepsilon(1) = 0$, we have

$$\int_0^1 \varepsilon|\nabla U(\gamma_\varepsilon)|\, dt = \int_0^1 \langle \varepsilon \nabla U(\gamma_\varepsilon), \nu(\gamma_\varepsilon)\rangle\, dt = \int_0^1 \langle \dot{\gamma}_\varepsilon, \nabla_t(\nu(\gamma_\varepsilon))\rangle\, dt.$$

Setting $M_0 := \max_{q \in Q} |\nabla \nu(q)|$, there holds

$$\int_0^1 \varepsilon|\nabla U(\gamma_\varepsilon)|\, dt \leq M_0 \|\dot{\gamma}_\varepsilon\|^2_{L^2} = 2M_0\Big(\mathcal{L}_\varepsilon(\gamma_\varepsilon) + \int_0^1 \varepsilon U(\gamma_\varepsilon)\, dt\Big). \tag{3.1}$$

By Lemma 2.1 (ii), there exists $M_1 > 0$ such that $U(q) \leq |\nabla U(q)|/4M_0 + M_1$ for any $q \in \operatorname{int} Q$. By the same arguments as in the proof of Lemma 2.7, we get

$$\frac{1}{2}\int_0^1 \varepsilon U(\gamma_\varepsilon)\, dt \leq \frac{1}{2}\mathcal{L}_\varepsilon(\gamma_\varepsilon) + \varepsilon M_1.$$

Since $\sup_\varepsilon \mathcal{L}_\varepsilon(\gamma_\varepsilon) \le b$, the above estimate implies $\sup_\varepsilon \int_0^1 \varepsilon U(\gamma_\varepsilon)\,dt < \infty$. By (3.1), we get $\sup_\varepsilon \int_0^1 \varepsilon|\nabla U(\gamma_\varepsilon)|\,dt < \infty$. The following identity is clear from the definition of U.

$$\int_0^1 \varepsilon|\nabla U(\gamma_\varepsilon)|\,dt = \int_0^1 2\varepsilon|\nabla h(\gamma_\varepsilon)|h(\gamma_\varepsilon)^{-3}\,dt.$$

Since $|\nabla h(q)| = 1$ for any q such that $h(q) < \delta$, we get $\sup_\varepsilon \int_0^1 \varepsilon h(\gamma_\varepsilon)^{-3}\,dt < \infty$. Finally, by the Hölder inequality, we obtain

$$\limsup_{\varepsilon\to 0} \int_0^1 \varepsilon h(\gamma_\varepsilon)^{-2}\,dt \le \limsup_{\varepsilon\to 0}\left(\int_0^1 \varepsilon h(\gamma_\varepsilon)^{-3}\,dt\right)^{2/3} \cdot \varepsilon^{1/3} = 0.$$

Since $0 \le \varepsilon U(q) \le \varepsilon h(q)^{-2}$ for any $q \in \mathrm{int}Q$, we obtain $\lim_{\varepsilon\to 0} \int_0^1 \varepsilon U(\gamma_\varepsilon)\,dt = 0$. □

Corollary 3.3. *The following quantities are bounded on* ε*.*

$$\int_0^1 |\ddot\gamma_\varepsilon(t)|\,dt = \int_0^1 \varepsilon|\nabla U(\gamma_\varepsilon)|\,dt, \quad \int_0^1 \varepsilon h(\gamma_\varepsilon)^{-3}\,dt, \quad E(\gamma_\varepsilon) := |\dot\gamma_\varepsilon|^2/2+\varepsilon U(\gamma_\varepsilon).$$

Proof. In the course of the proof of Lemma 3.2, we have shown that the first two quantities are bounded. $\sup_\varepsilon E(\gamma_\varepsilon) < \infty$ follows from the identity

$$E(\gamma_\varepsilon) = \int_0^1 \frac{|\dot\gamma_\varepsilon|^2}{2} + \varepsilon U(\gamma_\varepsilon)\,dt = \mathcal{L}_\varepsilon(\gamma_\varepsilon) + 2\int_0^1 \varepsilon U(\gamma_\varepsilon)\,dt$$

and estimates $\sup_\varepsilon \mathcal{L}_\varepsilon(\gamma_\varepsilon) \le b$, $\sup_\varepsilon \int_0^1 \varepsilon U(\gamma_\varepsilon)\,dt < \infty$. □

By Corollary 3.3, $\ddot\gamma_\varepsilon$ is L^1-bounded. Since $W^{2,1}([0,1])$ is compactly embedded to $W^{1,2}([0,1])$, a certain subsequence of $(\gamma_\varepsilon)_\varepsilon$ is convergent in $W^{1,2}([0,1],Q)$ as $\varepsilon \to 0$. We denote the limit as γ_0. Moreover, since $2\varepsilon h(\gamma_\varepsilon)^{-3}$ is L^1-bounded, up to subsequence it converges to a certain Borel measure $\mu \ge 0$ on $[0,1]$ in a weak sense, i.e., for any $f \in C^0([0,1])$ there holds

$$\lim_{\varepsilon\to 0} \int_0^1 f(t)2\varepsilon h(\gamma_\varepsilon(t))^{-3}\,dt = \int_0^1 f(t)\,d\mu(t).$$

For any $t \in [0,1]$ and $c > 0$, we set $B_c(t) := \{s \in [0,1] \mid |s-t| < c\}$. The *support* of μ is defined as $\mathrm{supp}\mu := \{t \in [0,1] \mid \forall c > 0, \mu(B_c(t)) > 0\}$.

Lemma 3.4. *There holds* $\mathrm{supp}\mu \subset \gamma_0^{-1}(\partial Q)$ *and* $\sharp\mathrm{supp}\mu \le j$.

Proof. If $\tau \in [0,1]$ satisfies $\gamma_0(\tau) \notin \partial Q$, $\varepsilon h(\gamma_\varepsilon(t))^{-3}$ converges uniformly to 0 in a neighborhood of τ, and thus, $\tau \notin \mathrm{supp}\mu$. Therefore, $\mathrm{supp}\mu \subset \gamma_0^{-1}(\partial Q)$.

We show that $\sharp\mathrm{supp}\mu \le j$. For any $\tau \in \mathrm{supp}\mu$, we have shown that $\gamma_0(\tau) \in \partial Q$. Hence, $d(\gamma_0(\tau)) = 0$. We take $c > 0$ so that $d(\gamma_0(t)) < \delta$ for any $t \in B_c(\tau)$.

We take $\psi \in C^\infty([0,1])$ so that $0 \le \psi(t) \le 1$ for any t, $\mathrm{supp}\psi \subset B_c(\tau)$, and $\psi \equiv 1$ on $B_{c/2}(\tau)$. Let $v_\varepsilon(t) := \psi(t)\nabla h(\gamma_\varepsilon(t))$. Our aim is to show that

$$\lim_{\varepsilon\to 0} d^2\mathcal{L}_\varepsilon(\gamma_\varepsilon)(v_\varepsilon, v_\varepsilon) = -\infty. \tag{3.2}$$

Obviously, $\mathrm{supp}v_\varepsilon \subset B_c(\tau)$, and we may take $c > 0$ arbitrarily small. Hence, once we prove (3.2), it is easy to show that $\liminf_{\varepsilon\to 0}\mathrm{ind}(\gamma_\varepsilon) \ge \sharp\mathrm{supp}\mu$. On the other hand, by our assumption in Proposition 3.1, $\mathrm{ind}(\gamma_\varepsilon) \le j$ for any $\varepsilon > 0$. Hence, $\sharp\mathrm{supp}\mu \le j$.

Now, we show (3.2). By (2.2), there holds

$$\begin{aligned} d^2\mathcal{L}_\varepsilon(\gamma_\varepsilon)(v_\varepsilon, v_\varepsilon) &= \int_0^1 |\nabla_t v_\varepsilon|^2 - \langle R(\dot\gamma_\varepsilon, v_\varepsilon)(v_\varepsilon), \dot\gamma_\varepsilon\rangle\, dt \\ &\quad + 2\varepsilon\int_0^1 \langle \nabla_{v_\varepsilon}\nabla h(\gamma_\varepsilon), v_\varepsilon\rangle h(\gamma_\varepsilon)^{-3}\, dt - 6\varepsilon\int_0^1 \{dh(\gamma_\varepsilon)(v_\varepsilon)\}^2 h(\gamma_\varepsilon)^{-4}\, dt. \end{aligned}$$

By Corollary 3.3, $\sup_\varepsilon \|\dot\gamma_\varepsilon\|_{L^\infty} < \infty$. Thus, it is easy to check that the first integral is bounded on ε. Corollary 3.3 also shows $\sup_\varepsilon \int_0^1 \varepsilon h(\gamma_\varepsilon)^{-3}\, dt < \infty$, and thus, the second integral is bounded on ε.

Recall that $d(\gamma_0(t)) < \delta$ for any $t \in B_c(\tau)$. Hence, when $\varepsilon > 0$ is sufficiently small, $d(\gamma_\varepsilon(t)) < \delta$ for any $t \in B_{c/2}(\tau)$. For such $\varepsilon > 0$, $dh(\gamma_\varepsilon)(v_\varepsilon) = |\nabla h(\gamma_\varepsilon)|^2 = 1$ on $B_{c/2}(\tau)$. Therefore,

$$\begin{aligned} \varepsilon\int_0^1 \{dh(\gamma_\varepsilon)(v_\varepsilon)\}^2 h(\gamma_\varepsilon)^{-4}\, dt &\ge \varepsilon\int_{B_{c/2}(\tau)} h(\gamma_\varepsilon)^{-4}\, dt \\ &\ge (c\varepsilon)^{-1/3}\Big(\int_{B_{c/2}(\tau)} \varepsilon h(\gamma_\varepsilon)^{-3}\, dt\Big)^{4/3}. \end{aligned}$$

The second inequality follows from the Hölder inequality. Since $\tau \in \mathrm{supp}\mu$,

$$\liminf_{\varepsilon\to 0}\int_{B_{c/2}(\tau)} \varepsilon h(\gamma_\varepsilon)^{-3}\, dt \ge \mu(B_{c/2}(\tau))/2 > 0.$$

Hence, $\lim_{\varepsilon\to 0}\varepsilon\int_0^1 \{dh(\gamma_\varepsilon)(v_\varepsilon)\}^2 h(\gamma_\varepsilon)^{-4}\, dt = \infty$, and therefore, we have proved (3.2). □

For $q \in \partial Q$, let $\nu(q)$ denote the unit vector that is outer normal to ∂Q at q.

Lemma 3.5. *For any $v \in W^{1,2}([0,1], \gamma_0^*(TQ))$, there holds*

$$\int_0^1 \langle \dot\gamma_0, \nabla_t v\rangle\, dt = \int_0^1 \langle \nu(\gamma_0), v\rangle\, d\mu(t).$$

Notice that the RHS is well-defined, since $\mathrm{supp}\mu \subset \gamma_0^{-1}(\partial Q)$.

Proof. One can take $v_\varepsilon \in T_{\gamma_\varepsilon}\Lambda$ so that $v_\varepsilon \to v$ as $\varepsilon \to 0$, in $W^{1,2}$-norm. By $\ddot{\gamma}_\varepsilon + \varepsilon\nabla U(\gamma_\varepsilon) \equiv 0$ and $\dot{\gamma}_\varepsilon(0) = \dot{\gamma}_\varepsilon(1) = 0$, we get

$$\int_0^1 \langle \varepsilon\nabla U(\gamma_\varepsilon), v_\varepsilon(t)\rangle\, dt = -\int_0^1 \langle \ddot{\gamma}_\varepsilon(t), v_\varepsilon(t)\rangle\, dt = \int_0^1 \langle \dot{\gamma}_\varepsilon(t), \nabla_t(v_\varepsilon(t))\rangle\, dt.$$

As $\varepsilon \to 0$, RHS goes to $\int_0^1 \langle \dot{\gamma}_0, \nabla_t v\rangle\, dt$. On the other hand, since $\nabla U(q) = -2\nabla h(q)h(q)^{-3}$, LHS goes to $\int_0^1 \langle \nu(\gamma_0), v\rangle\, d\mu(t)$ as $\varepsilon \to 0$. □

Lemma 3.5 shows that $\ddot{\gamma}_0 \equiv 0$ on $[0,1] \setminus \operatorname{supp}\mu$. Lemma 3.4 shows that $\operatorname{supp}\mu$ is discrete. Hence, $\dot{\gamma}_0^-(t) = \lim_{h\to 0-} \dot{\gamma}_0(t+h)$ exists for any $t > 0$, and $\dot{\gamma}_0^+(t) = \lim_{h\to 0+} \dot{\gamma}_0(t+h)$ exists for any $t < 1$. Now, we show that γ_0 satisfies the following properties:

- $\operatorname{length}(\gamma_0) \in [\sqrt{2a}, \sqrt{2b}]$.
- $\{0,1\} \subset \operatorname{supp}\mu$. Moreover, $\dot{\gamma}_0^+(0)$, $\dot{\gamma}_0^-(1)$ are perpendicular to ∂Q.
- γ_0 satisfies the law of reflection at every point on $\operatorname{supp}\mu \setminus \{0,1\}$.

Once these properties are confirmed, γ_0 is a brake billiard trajectory with $\mathcal{B}_{\gamma_0} = \operatorname{supp}\mu \setminus \{0,1\}$, and Proposition 3.1 (i) is proved.

Let I be any interval on $[0,1]$. By Lemma 3.2,

$$\begin{aligned}\int_I |\dot{\gamma}_0|^2\, dt = \lim_{\varepsilon\to 0}\int_I |\dot{\gamma}_\varepsilon|^2\, dt &= \lim_{\varepsilon\to 0} 2\Big(|I|E(\gamma_\varepsilon) - \int_I \varepsilon U(\gamma_\varepsilon)\, dt\Big)\\ &= 2|I| \lim_{\varepsilon\to 0} E(\gamma_\varepsilon).\end{aligned}$$

Hence, $E := \lim_{\varepsilon\to 0} E(\gamma_\varepsilon)$ exists, and $|\dot{\gamma}_0(t)| = \sqrt{2E}$ holds for any $t \notin \operatorname{supp}\mu$. Then, $\operatorname{length}(\gamma_0) \in [\sqrt{2a}, \sqrt{2b}]$ follows from

$$E = \lim_{\varepsilon\to 0} E(\gamma_\varepsilon) = \lim_{\varepsilon\to 0} \mathcal{L}_\varepsilon(\gamma_\varepsilon) + 2\int_0^1 \varepsilon U(\gamma_\varepsilon)\, dt = \lim_{\varepsilon\to 0} \mathcal{L}_\varepsilon(\gamma_\varepsilon) \in [a,b].$$

Let us prove that $0 \in \operatorname{supp}\mu$. If not, there exists $c > 0$ such that $\mu \equiv 0$ on $[0,c]$. Take $f \in C^\infty([0,1])$ such that $f(0) = 1$ and $\operatorname{supp} f \subset [0,c]$. Let $v(t) := f(t)\dot{\gamma}_0(t)$. Then, Lemma 3.5 implies

$$0 = \int_0^1 \langle \dot{\gamma}_0, \nabla_t v\rangle\, dt = \int_0^1 f'(t)|\dot{\gamma}_0(t)|^2\, dt = -2E.$$

This contradicts $E \in [a,b]$ and $a > 0$, hence $0 \in \operatorname{supp}\mu$. We can show that $1 \in \operatorname{supp}\mu$ by the same arguments.

Let us prove that $\dot{\gamma}_0^+(0)$ is perpendicular to ∂Q. Let ζ_0 be any tangent vector of ∂Q at $\gamma_0(0)$. Take $c > 0$ sufficiently small so that $[0,c] \cap \operatorname{supp}\mu = \{0\}$, and define

$\zeta(t) \in T_{\gamma_0(t)}Q$ for any $0 \le t \le c$ by $\zeta(0) = \zeta_0$ and $\nabla_t \zeta \equiv 0$. Take $f \in C^\infty([0,1])$ as above, and set $v(t) := f(t)\zeta(t)$. Then, Lemma 3.5 implies

$$\mu(\{0\})\langle \nu(\gamma_0(0)), \zeta_0 \rangle = \int_0^1 \langle \nu(\gamma_0), v \rangle \, d\mu(t) = \int_0^1 \langle \dot{\gamma}_0, \nabla_t v \rangle dt = -\langle \zeta_0, \dot{\gamma}_0^+(0) \rangle.$$

Since ζ_0 is tangent to ∂Q, LHS is zero, and therefore, $\langle \zeta_0, \dot{\gamma}_0^+(0) \rangle = 0$. This shows that $\dot{\gamma}_0^+(0)$ is perpendicular to ∂Q. By the same arguments, we can show that $\dot{\gamma}_0^-(1)$ is perpendicular to ∂Q.

Finally, let us prove that γ_0 satisfies the law of reflection at any $t \in \mathrm{supp}\mu \setminus \{0,1\}$. Similar arguments as above show that $\dot{\gamma}_0^+(t) - \dot{\gamma}_0^-(t)$ is nonzero and perpendicular to ∂Q. On the other hand, $|\dot{\gamma}_0^+(t)| = |\dot{\gamma}_0^-(t)|$, since both are equal to $\sqrt{2E}$. Then, it is immediate that γ_0 satisfies the law of reflection at t.

We have now finished the proof of Proposition 3.1 (i). As we explained at the beginning of this section, (ii) can be proved by parallel arguments.

4. Proof of Theorem 1.2

In this section, we complete the proof of Theorem 1.2. We only prove (i), since (ii) can be proved by parallel arguments. We may assume that Q is connected and $\partial Q \neq \emptyset$ (see Remark 1.3). First, we need the following technical lemma. Let us denote

$$\Lambda^c(\mathrm{int}Q) := \Lambda^c(Q) \cap \Lambda(\mathrm{int}Q), \quad \Lambda_\delta(\mathrm{int}Q) := \Lambda_\delta(Q) \cap \Lambda(\mathrm{int}Q).$$

Lemma 4.1. *For any $c \in \mathbb{R}$ and $\delta > 0$, there holds*

$$H_*(\Lambda^c(Q) \cup \Lambda_\delta(Q), \Lambda^c(\mathrm{int}Q) \cup \Lambda_\delta(\mathrm{int}Q)) = 0.$$

Proof. It is sufficient to show that the inclusion

$$\Lambda^c(\mathrm{int}Q) \cup \Lambda_\delta(\mathrm{int}Q) \to \Lambda^c(Q) \cup \Lambda_\delta(Q) \tag{4.1}$$

is a homotopy equivalence. Let Z be a smooth vector field on Q, which points strictly inwards on ∂Q, and $Z \equiv 0$ on $Q(\delta)$. Let $(\psi^t)_{t \ge 0}$ be the isotopy generated by Z, i.e., it satisfies $\psi^0 = \mathrm{id}_Q$ and $\partial_t \psi^t = Z(\psi^t)$. Then, it is easy to show that

$$\Lambda^c(Q) \cup \Lambda_\delta(Q) \to \Lambda^c(\mathrm{int}Q) \cup \Lambda_\delta(\mathrm{int}Q); \quad \gamma \mapsto \psi^1 \circ \gamma$$

is a homotopy inverse of (4.1). □

By Lemma 4.1, the assumption of Theorem 1.2 (i) is equivalent to

$$\varprojlim_{\delta \to 0} H_j(\Lambda^b(\mathrm{int}Q) \cup \Lambda_\delta(\mathrm{int}Q), \Lambda^a(\mathrm{int}Q) \cup \Lambda_\delta(\mathrm{int}Q)) \neq 0.$$

In this section, we abbreviate $\Lambda^b(\mathrm{int} Q)$ as Λ^b, $\Lambda_\delta(\mathrm{int} Q)$ as Λ_δ, and so on. There exists $\delta_0 > 0$ such that

$$\varinjlim_{\delta\to 0} H_j(\Lambda^b \cup \Lambda_\delta, \Lambda^a \cup \Lambda_\delta) \to H_j(\Lambda^b \cup \Lambda_{\delta_0}, \Lambda^a \cup \Lambda_{\delta_0}) \tag{4.2}$$

is nonzero. We take $\delta_1 > 0$ so that $3\delta_1 \le \delta_0$. We are going to prove

$$H_j(\{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < b\}, \{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < a\}) \neq 0$$

for any $\varepsilon > 0$. Once we prove this, Proposition 2.2 and Proposition 3.1 show that there exists a brake billiard trajectory γ such that $\sharp\mathcal{B}_\gamma \le j-2$ and length$(\gamma) \in [\sqrt{2a}, \sqrt{2b}]$.

We fix $\varepsilon > 0$. For any $c > 0$, there holds $\{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < c\} \subset \Lambda^c \cup \Lambda_{\delta_0}$ since $U_{\delta_1} \equiv 0$ on $Q(\delta_0)$. On the other hand, Lemma 2.3 shows that, for sufficiently small $\delta_2 > 0$, there holds $\Lambda^b \cap \Lambda_{\delta_2} \subset \{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < a\}$. Thus, we have the following commutative diagram, where all homomorphisms are induced by inclusions.

$$\begin{array}{ccc} H_j(\Lambda^b, \Lambda^a \cup (\Lambda^b \cap \Lambda_{\delta_2})) & \longrightarrow & H_j(\{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < b\}, \{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < a\}) \\ \downarrow & & \downarrow \\ H_j(\Lambda^b \cup \Lambda_{\delta_2}, \Lambda^a \cup \Lambda_{\delta_2}) & \longrightarrow & H_j(\Lambda^b \cup \Lambda_{\delta_0}, \Lambda^a \cup \Lambda_{\delta_0}). \end{array}$$

Since (4.2) is nonzero, the bottom arrow is nonzero. On the other hand, the excision property shows that the left vertical arrow is an isomorphism. By commutativity of the diagram, we have $H_j(\{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < b\}, \{\mathcal{L}^\Lambda_{\varepsilon U_{\delta_1}} < a\}) \neq 0$, and this completes the proof.

5. Short billiard trajectory

In this section, we prove Theorem 1.4. In Section 5.1, we introduce the notion of capacity for Riemannian manifolds with boundaries, and show that the capacity is equal to the length of a billiard trajectory (Lemma 5.4). In Section 5.2, we bound the capacity by the inradius, and complete the proof of Theorem 1.4. In Section 5.3, we prove Corollary 1.7 as a consequence of Theorem 1.4.

5.1. Capacity.

First, we introduce some notations.

- We define $\Lambda_\partial(Q) \subset \Lambda(Q)$, $\Omega_\partial(Q) \subset \Omega(Q)$ as

$$\Lambda_\partial(Q) := \{\gamma \in \Lambda(Q) \mid \gamma([0,1]) \cap \partial Q \neq \emptyset\},$$
$$\Omega_\partial(Q) := \{\gamma \in \Omega(Q) \mid \gamma(S^1) \cap \partial Q \neq \emptyset\}.$$

- For each $q \in Q$, p_q denotes the constant path at q, and l_q denotes the constant loop at q.

We often identify $q \in Q$ with p_q and l_q, and thus, we have inclusions $Q \to \Lambda(Q)$, $Q \to \Omega(Q)$. For each $a > 0$, we consider the following homomorphisms, all induced by inclusions.

$$\begin{aligned}
I_0^{\Lambda,a} &: H_*(Q, \partial Q) \to H_*(\Lambda^a(Q) \cup \Lambda_\partial(Q), \Lambda_\partial(Q)),\\
I_1^{\Lambda,a} &: H_*(Q, \partial Q) \to \varprojlim_{\delta \to 0} H_*(\Lambda^a(Q) \cup \Lambda_\delta(Q), \Lambda_\delta(Q)),\\
I_2^{\Lambda,a} &: H_*(Q, \partial Q) \cong \varprojlim_{\delta \to 0} H_*(\mathrm{int}Q, \mathrm{int}Q \setminus Q(\delta))\\
&\qquad\qquad \to \varprojlim_{\delta \to 0} H_*(\Lambda^a(\mathrm{int}Q) \cup \Lambda_\delta(\mathrm{int}Q), \Lambda_\delta(\mathrm{int}Q)).
\end{aligned}$$

One can define $I_0^{\Omega,a}$, $I_1^{\Omega,a}$, and $I_2^{\Omega,a}$ in the same manner.

Lemma 5.1. *For any $\alpha \in H_*(Q, \partial Q)$ and $j = 0, 1, 2$, let us define*

$$c_j^\Lambda(\alpha) := \inf\{c > 0 \mid I_j^{\Lambda, c^2/2}(\alpha) = 0\}.$$

Then, $c_0^\Lambda(\alpha) = c_1^\Lambda(\alpha) = c_2^\Lambda(\alpha)$.

Proof. $c_1^\Lambda(\alpha) = c_2^\Lambda(\alpha)$ is immediate from Lemma 4.1. $c_1^\Lambda(\alpha) \le c_0^\Lambda(\alpha)$ is also clear, since there exists a natural homomorphism

$$H_*(\Lambda^a(Q) \cup \Lambda_\partial(Q), \Lambda_\partial(Q)) \to \varprojlim_{\delta \to 0} H_*(\Lambda^a(Q) \cup \Lambda_\delta(Q), \Lambda_\delta(Q)),$$

which is induced by inclusions. Hence, it is sufficient to prove $c_1^\Lambda(\alpha) \ge c_0^\Lambda(\alpha)$.

Let $a > a'$ be any positive real numbers. When $\delta > 0$ is sufficiently small, there exists a C^∞ map $\psi : Q \times [0,1] \to Q; (x,t) \mapsto \psi_t(x)$ such that

- $\psi_0 = \mathrm{id}_Q$. $\psi_t|_{\partial Q} = \mathrm{id}_{\partial Q}$ for any $0 \le t \le 1$.
- $\psi_1(Q \setminus Q(\delta)) = \partial Q$.
- $|d\psi_1(\xi)| \le \sqrt{a/a'}|\xi|$ for any $\xi \in TQ$.

Then, we have the commutative diagram

$$\begin{array}{ccc}
H_*(Q, \partial Q) & \xrightarrow{\qquad\qquad\qquad} & H_*(\Lambda^{a'}(Q) \cup \Lambda_\delta(Q), \Lambda_\delta(Q)) \\
& {\scriptstyle I_0^{\Lambda,a}} \searrow \qquad \swarrow {\scriptstyle (\psi_1)_*} & \\
& H_*(\Lambda^a(Q) \cup \Lambda_\partial(Q), \Lambda_\partial(Q)). &
\end{array}$$

If $a' > c_1^\Lambda(\alpha)^2/2$, $\alpha \in H_*(Q, \partial Q)$ vanishes by the top arrow, hence $I_0^{\Lambda,a}(\alpha) = 0$, and therefore, $a > c_0^\Lambda(\alpha)^2/2$. Since we may take $a > a'$ arbitrarily, we have shown that $c_1^\Lambda(\alpha) \ge c_0^\Lambda(\alpha)$. □

For any $\alpha \in H_*(Q, \partial Q)$, we denote $c_0^\Lambda(\alpha) = c_1^\Lambda(\alpha) = c_2^\Lambda(\alpha)$ in Lemma 5.1 by $c^\Lambda(Q : \alpha)$. On the other hand, for $j = 0, 1, 2$, we define

$$c_j^\Omega(\alpha) := \inf\{c > 0 \mid I_j^{\Omega, c^2/2}(\alpha) = 0\}.$$

By the same arguments as in Lemma 5.1, we can show that $c_0^\Omega(\alpha) = c_1^\Omega(\alpha) = c_2^\Omega(\alpha)$. We denote it by $c^\Omega(Q : \alpha)$. We call $c^\Lambda(Q : \alpha)$ and $c^\Omega(Q : \alpha)$ the *capacities* of Q.

Remark 5.2. The above definition of c^Λ and c^Ω imitate the definition of the Floer–Hofer–Wysocki (FHW) capacity, which is defined in [9] (see also [12], Section 2.4). In fact, when Q is a domain in the Euclidean space and $[Q, \partial Q]$ denotes its relative fundamental class, $c^\Omega(Q : [Q, \partial Q])$ is equal to the FHW capacity of its disc cotangent bundle. See Corollary 1.4 in [12].

Lemma 5.3. *For any* $\alpha \in H_*(Q, \partial Q) \setminus \{0\}$, $c^\Lambda(Q : \alpha), c^\Omega(Q : \alpha) > 0$.

Proof. We only prove $c^\Lambda(Q : \alpha) > 0$, since $c^\Omega(Q : \alpha) > 0$ can be proved by parallel arguments. In this proof, we use abbreviations $\Lambda^a := \Lambda^a(\mathrm{int}Q)$, $\Lambda_\delta := \Lambda_\delta(\mathrm{int}Q)$. For any positive a and δ, the excision property shows that

$$H_*(\Lambda^a, \Lambda^a \cap \Lambda_\delta) \to H_*(\Lambda^a \cup \Lambda_\delta, \Lambda_\delta)$$

is an isomorphism. Therefore, it is sufficient to show that for sufficiently small $a > 0$

$$\varprojlim_{\delta \to 0} H_*(\mathrm{int}Q, \mathrm{int}Q \setminus Q(\delta)) \to \varprojlim_{\delta \to 0} H_*(\Lambda^a, \Lambda^a \cap \Lambda_\delta)$$

is injective. For any $\gamma \in \Lambda^a \cap \Lambda_\delta$, there holds

$$\gamma(0) \in \mathrm{int}Q \setminus Q(\delta + \mathrm{length}(\gamma)) \subset \mathrm{int}Q \setminus Q(\delta + \sqrt{2a}).$$

Define $\mathrm{ev} : \Lambda^a \to \mathrm{int}Q$ by $\mathrm{ev}(\gamma) := \gamma(0)$, and consider the commutative diagram

$$\begin{array}{ccc} \varprojlim_{\delta \to 0} H_*(\mathrm{int}Q, \mathrm{int}Q \setminus Q(\delta)) & \longrightarrow & \varprojlim_{\delta \to 0} H_*(\Lambda^a, \Lambda^a \cap \Lambda_\delta) \\ & \searrow & \downarrow (\mathrm{ev})_* \\ & & \varprojlim_{\delta \to 0} H_*(\mathrm{int}Q, \mathrm{int}Q \setminus Q(\delta + \sqrt{2a})). \end{array}$$

When $a > 0$ is sufficiently small, the diagonal arrow is an isomorphism. Therefore, the horizontal arrow is injective. □

The next lemma shows that the capacity is equal to the length of a billiard trajectory.

Lemma 5.4. *Let* $\alpha \in H_j(Q, \partial Q) \setminus \{0\}$.

(i): *If* $c^{\Lambda}(Q : \alpha) < \infty$, *there exists a brake billiard trajectory* γ *on* Q *such that* $\sharp\mathcal{B}_\gamma \le j-1$ *and* $\mathrm{length}(\gamma) = c^{\Lambda}(Q : \alpha)$.

(ii): *If* $c^{\Omega}(Q : \alpha) < \infty$, *there exists a periodic billiard trajectory* γ *on* Q *such that* $\sharp\mathcal{B}_\gamma \le j+1$ *and* $\mathrm{length}(\gamma) = c^{\Omega}(Q : \alpha)$.

Proof. We only prove (i), since (ii) can be proved by parallel arguments. We set $a := c^{\Lambda}(Q : \alpha)^2/2$. Then, for any $\varepsilon > 0$, there holds $I_1^{\Lambda, a-\varepsilon}(\alpha) \neq 0$ and $I_0^{\Lambda, a+\varepsilon}(\alpha) = 0$. In this proof, we use the notations $\Lambda^a := \Lambda^a(Q)$, $\Lambda_\delta := \Lambda_\delta(Q)$, $\Lambda_\partial := \Lambda_\partial(Q)$, and so on.

For any $\delta > 0$, we have a commutative diagram

$$\begin{array}{ccc} H_{j+1}(\Lambda^{a+\varepsilon} \cup \Lambda_\partial, \Lambda^{a-\varepsilon} \cup \Lambda_\partial) & \xrightarrow{\partial_0} & H_j(\Lambda^{a-\varepsilon} \cup \Lambda_\partial, \Lambda_\partial) \\ \downarrow & & \downarrow \\ H_{j+1}(\Lambda^{a+\varepsilon} \cup \Lambda_\delta, \Lambda^{a-\varepsilon} \cup \Lambda_\delta) & \xrightarrow[\partial_\delta]{} & H_j(\Lambda^{a-\varepsilon} \cup \Lambda_\delta, \Lambda_\delta), \end{array}$$

where vertical arrows are induced by inclusions, and horizontal arrows are connecting homomorphisms. Since $I_0^{\Lambda, a+\varepsilon}(\alpha) = 0$, we have $I_0^{\Lambda, a-\varepsilon}(\alpha) \in \mathrm{Im}\partial_0$. Letting $\delta \to 0$ of the above diagram, we have the following commutative diagram.

$$\begin{array}{ccc} H_{j+1}(\Lambda^{a+\varepsilon} \cup \Lambda_\partial, \Lambda^{a-\varepsilon} \cup \Lambda_\partial) & \xrightarrow{\partial_0} & H_j(\Lambda^{a-\varepsilon} \cup \Lambda_\partial, \Lambda_\partial) \\ \downarrow & & \downarrow \iota \\ \varprojlim_{\delta \to 0} H_{j+1}(\Lambda^{a+\varepsilon} \cup \Lambda_\delta, \Lambda^{a-\varepsilon} \cup \Lambda_\delta) & \longrightarrow & \varprojlim_{\delta \to 0} H_j(\Lambda^{a-\varepsilon} \cup \Lambda_\delta, \Lambda_\delta). \end{array}$$

Let us denote the right vertical arrow as ι. Then, $\iota(I_0^{\Lambda, a-\varepsilon}(\alpha)) = I_1^{\Lambda, a-\varepsilon}(\alpha) \neq 0$. Since $I_0^{\Lambda, a-\varepsilon}(\alpha) \in \mathrm{Im}\partial_0$, we get $\varprojlim_{\delta \to 0} H_{j+1}(\Lambda^{a+\varepsilon} \cup \Lambda_\delta, \Lambda^{a-\varepsilon} \cup \Lambda_\delta) \neq 0$. By Theorem 1.2, there exists a brake billiard trajectory γ_ε on Q such that $\sharp\mathcal{B}_{\gamma_\varepsilon} \le j-1$ and $\mathrm{length}(\gamma_\varepsilon) \in [\sqrt{2(a-\varepsilon)}, \sqrt{2(a+\varepsilon)}]$. As $\varepsilon \to 0$, a certain subsequence of $(\gamma_\varepsilon)_\varepsilon$ converges to a brake billiard trajectory γ such that $\sharp\mathcal{B}_\gamma \le j-1$ and $\mathrm{length}(\gamma) = \sqrt{2a} = c^{\Lambda}(Q : \alpha)$. □

5.2. Capacity and inradius. By Lemma 5.4, Theorem 1.4 follows at once from the following proposition. Recall that $r(Q)$ denotes the inradius of Q.

Proposition 5.5. *Let* Q *be a compact, connected Riemannian manifold with nonempty boundary, and* $\alpha \in H_j(Q, \partial Q)$. *Then, there holds* $c^{\Lambda}(Q : \alpha) \le 2jr(Q)$, $c^{\Omega}(Q : \alpha) \le 2(j+1)r(Q)$.

Proposition 5.5 is proved in this subsection. We will give a proof that stems from arguments in our paper [12], Section 7. First, we need some preliminary results: Lemma 5.6, Lemma 5.7.

Let P be a finite simplicial complex and σ be a simplex on P. $\mathrm{Star}(\sigma) \subset P$ denotes the union of interiors of all simplices of P which contain σ as a facet, i.e., $\mathrm{Star}(\sigma) := \bigcup_{\sigma \subset \tau} \mathrm{int}\tau$.

Lemma 5.6. *Let P be a finite simplicial complex. There exist continuous functions $w_\sigma : P \to [0,1]$ where σ runs over all simplices of P, such that the following holds.*

(i): *For any simplex σ,* $\mathrm{supp} w_\sigma \subset \mathrm{Star}(\sigma)$.

(ii): *For any distinct simplices σ, σ' of same dimensions,* $\mathrm{supp} w_\sigma \cap \mathrm{supp} w_{\sigma'} = \emptyset$.

(iii): $\bigcup_\sigma w_\sigma^{-1}(1) = P$, *where σ runs over all simplices of P.*

Proof. We prove the lemma by induction on $\dim P$. The claim is obvious when $\dim P = 0$. Suppose that we have proved the claim for finite simplicial complexes of dimension $\le d-1$, and let P be a finite simplicial complex of dimension d.

Let $\sigma_1, \dots, \sigma_m$ be all simplices on P of dimension d, and $P^{(d-1)}$ denote the union of all simplices on P of dimension $\le d-1$. Take $x_j \in \mathrm{int}\sigma_j$ for every $j = 1, \dots, m$. There exists a continuous retraction $r : P \setminus \{x_1, \dots, x_m\} \to P^{(d-1)}$ such that there holds $r(\sigma_j \setminus \{x_j\}) = \partial\sigma_j$ for any $j = 1, \dots, m$.

We define a continuous function $\tilde{w}_\sigma : P \to [0,1]$ for each simplex σ of P. When $\dim\sigma = d$, i.e., $\sigma = \sigma_j$ for some $j = 1, \dots, m$, we define $\tilde{w}_{\sigma_j}$ so that $\mathrm{supp}\tilde{w}_{\sigma_j} \subset \mathrm{int}\sigma_j$, and $\tilde{w}_{\sigma_j} \equiv 1$ on some neighborhood of x_j. Then, there exists a continuous function $v : P \to [0,1]$ such that $x_1, \dots, x_m \notin \mathrm{supp} v$ and $\tilde{w}_{\sigma_1}^{-1}(1) \cup \dots \cup \tilde{w}_{\sigma_m}^{-1}(1) \cup v^{-1}(1) = P$.

Next, we define $\tilde{w}_\sigma$ when $\dim\sigma \le d-1$. By induction hypothesis, one can take $w_\sigma : P^{(d-1)} \to [0,1]$ for each $\sigma \subset P^{(d-1)}$ so that our requirements (i)–(iii) hold for $(w_\sigma)_{\sigma \subset P^{(d-1)}}$. We define $\tilde{w}_\sigma : P \to [0,1]$ by

$$\tilde{w}_\sigma(x) := \begin{cases} 0 & (x \in \{x_1, \dots, x_m\}), \\ v(x) w_\sigma(r(x)) & (x \notin \{x_1, \dots, x_m\}). \end{cases}$$

Let us check that $(\tilde{w}_\sigma)_\sigma$ satisfies our requirements (i)–(iii). By definition, if $\dim\sigma = d$, then $\mathrm{supp}\tilde{w}_\sigma \subset \mathrm{int}\sigma$. Then, (i), (ii) are obvious when $\dim\sigma = d$. If $\dim\sigma \le d-1$, then $\mathrm{supp}\tilde{w}_\sigma \subset r^{-1}(\mathrm{supp} w_\sigma)$. This is because $\{\tilde{w}_\sigma \ne 0\}$ is contained in $\mathrm{supp} v \cap r^{-1}(\mathrm{supp} w_\sigma)$, which is closed in P. Then, one can prove (i) for $\dim\sigma \le d-1$ by

$$\mathrm{supp}\tilde{w}_\sigma \subset r^{-1}(\mathrm{supp} w_\sigma) \subset r^{-1}(\mathrm{Star}(\sigma) \cap P^{(d-1)}) \subset \mathrm{Star}(\sigma).$$

The second inclusion holds since $(w_\sigma)_\sigma$ satisfies (i), and the third inclusion holds since $r(\sigma_j \setminus \{x_j\}) = \partial\sigma_j$ for any $j = 1, \dots, m$. (ii) for $\dim\sigma \le d-1$ is proved as follows (notice that $\mathrm{supp} w_\sigma \cap \mathrm{supp} w_{\sigma'} = \emptyset$, since $(w_\sigma)_\sigma$ satisfies (ii)):

$$\mathrm{supp}\tilde{w}_\sigma \cap \mathrm{supp}\tilde{w}_{\sigma'} \subset r^{-1}(\mathrm{supp} w_\sigma \cap \mathrm{supp} w_{\sigma'}) = \emptyset.$$

(iii) follows from $\bigcup_{\sigma \subset P^{(d-1)}} w_\sigma^{-1}(1) = P^{(d-1)}$ (since $(w_\sigma)_\sigma$ satisfies (iii)) and $\tilde{w}_{\sigma_1}^{-1}(1) \cup \cdots \cup \tilde{w}_{\sigma_m}^{-1}(1) \cup v^{-1}(1) = P$. □

Lemma 5.7. *For any $R > r(Q)^2/2$ and $q \in Q$, there exists an open neighborhood V of q and a continuous map $\lambda : V \to \Lambda^R(Q)$ such that there holds $\lambda(v)(0) = v$ and $\lambda(v)(1) \in \partial Q$ for any $v \in V$.*

Proof. Since $R > r(Q)^2/2 \geq \mathrm{dist}(q, \partial Q)^2/2$, there exists $\gamma \in \Lambda^R(Q)$ such that $\gamma(0) = q$ and $\gamma(1) \in \partial Q$. Then, there exists an open neighborhood $\tilde{V}$ of q and a continuous map $\tilde{\lambda} : \tilde{V} \to \Lambda(Q)$ such that $\tilde{\lambda}(q) = \gamma$ and $\tilde{\lambda}(v)(0) = v$, $\tilde{\lambda}(v)(1) \in \partial Q$ ($\forall v \in \tilde{V}$). Then, $V := \tilde{\lambda}^{-1}(\Lambda^R(Q))$ and $\lambda := \tilde{\lambda}|_V$ satisfy our requirements. □

Before starting the proof of Proposition 5.5, we introduce some operations on $\Lambda(Q)$.

- For any $a \in [0, 1]$ and $\gamma \in \Lambda(Q)$, we define $a\gamma \in \Lambda(Q)$ by $a\gamma(t) := \gamma(at)$. The map $[0, 1] \times \Lambda(Q) \to \Lambda(Q); (a, \gamma) \mapsto a\gamma$ is continuous.
- For any $\gamma \in \Lambda(Q)$, we define $\bar{\gamma} \in \Lambda(Q)$ by $\bar{\gamma}(t) := \gamma(1 - t)$. The map $\Lambda(Q) \to \Lambda(Q); \gamma \mapsto \bar{\gamma}$ is continuous.
- For any $\gamma_1, \ldots, \gamma_m \in \Lambda(Q)$ such that $\gamma_k(1) = \gamma_{k+1}(0)$ for $k = 1, \ldots, m-1$, We define $\mathrm{con}(\gamma_1, \ldots, \gamma_m) \in \Lambda(Q)$ by

  $$\mathrm{con}(\gamma_1, \ldots, \gamma_m)(t) := \gamma_{k+1}(m(t - k/m)) \quad (k/m \leq t \leq (k+1)/m,\ k = 0, \ldots, m-1).$$

  This is called the *concatenation* of $\gamma_1, \ldots, \gamma_m$. The following map is continuous:

  $$\{(\gamma_1, \ldots, \gamma_m) \mid \gamma_1, \ldots, \gamma_m \in \Lambda(Q), \gamma_k(1) = \gamma_{k+1}(0)\ (k = 1, \ldots, m-1)\} \to \Lambda(Q); (\gamma_1, \ldots, \gamma_m) \mapsto \mathrm{con}(\gamma_1, \ldots, \gamma_m).$$

Proof of Proposition 5.5. First, we prove $c^\Lambda(Q : \alpha) \leq 2jr(Q)$. It is sufficient to show $I_0^{\Lambda,a}(\alpha) = 0$ for any $a > (2jr(Q))^2/2$. Let us take a j-dimensional finite simplicial complex P, a subcomplex $P' \subset P$, and a continuous map $f : (P, P') \to (Q, \partial Q)$ such that $\alpha \in f_*(H_j(P, P'))$.

Suppose that there exists a continuous map $F : P \times [0,1] \to \Lambda^a(Q)$ that satisfies the following properties.

F-(i): For any $x \in P$, $F(x,0) = p_{f(x)}$.

F-(ii): For any $(x,t) \in P'' := P' \times [0,1] \cup P \times \{1\}$, $F(x,t) \in \Lambda_\partial(Q)$.

We obtain the following commutative diagram, where $i^P : (P, P') \to (P \times [0,1], P'')$ is defined by $i^P(x) := (x, 0)$.

$$\begin{array}{ccccc} H_j(P,P') & \xrightarrow{f_*} & H_j(Q,\partial Q) & \xrightarrow{I_0^{\Lambda,a}} & \\ \downarrow (i^P)_* & & \downarrow & & \\ H_j(P\times[0,1],P'') & \xrightarrow[F_*]{} & H_j(\Lambda^a(Q),\Lambda^a(Q)\cap\Lambda_\partial(Q)) & \longrightarrow & H_j(\Lambda^a(Q)\cup\Lambda_\partial(Q),\Lambda_\partial(Q)). \end{array}$$

It is easy to see that $(i^P)_* = 0$, thus $I_0^{\Lambda,a} \circ f_* = 0$. Since $\alpha \in f_*(H_j(P,P'))$, we have $I_0^{\Lambda,a}(\alpha) = 0$. Hence, it is sufficient to define F that satisfies F-(i) and F-(ii).

By our assumption, $a/(2j)^2 > r(Q)^2/2$. By Lemma 5.7, for any $q \in Q$ there exists a neighborhood V_q of q and $\lambda_q : V_q \to \Lambda^{a/(2j)^2}(Q)$ that satisfies $\lambda_q(v)(0) = v$ and $\lambda_q(v)(1) \in \partial Q$ for any $v \in V_q$.

By replacing P with its subdivison if necessary, we may assume that the following holds: for any simplex σ of P, there exists $q \in Q$ such that $f(\mathrm{Star}(\sigma)) \subset V_q$. We choose such q, and denote it by $q(\sigma)$. Moreover, we take $(w_\sigma)_\sigma$, a family of continuous functions on P as in Lemma 5.6.

We define $F_k : P \to \Lambda(Q)$ for each $k = 0, \dots, j$. Since $(w_\sigma)_\sigma$ satisfies Lemma 5.6 (ii), for each $x \in P$ and $k = 0, \dots, j$, either (a) or (b) holds.

(a): There exists a unique k-dimensional simplex σ of P such that $x \in \mathrm{supp} w_\sigma$.

(b): $x \notin \mathrm{supp} w_\sigma$ for any k-dimensional simplex σ of P.

In case (a), $f(x) \in f(\mathrm{Star}(\sigma)) \subset V_{q(\sigma)}$. Then, we define $F_k(x) \in \Lambda(Q)$ by $F_k(x) := w_\sigma(x) \cdot \lambda_{q(\sigma)}(f(x))$, i.e.,

$$F_k(x) : [0,1] \to Q; \quad t \mapsto \lambda_{q(\sigma)}(f(x))(w_\sigma(x) \cdot t).$$

In case (b), we define $F_k(x) := p_{f(x)}$. Then, it is easy to check that F_k is a continuous map, which satisfies the following properties.

- For any $x \in P$, $\mathcal{E}(F_k(x)) < a/(2j)^2$.
- For any $x \in P$, $F_k(x)(0) = f(x)$.
- If $x \in P$ satisfies $w_\sigma(x) = 1$ for some k-dimensional simplex σ of P, $F_k(x)(1) = \lambda_{q(\sigma)}(f(x))(1) \in \partial Q$.

Now, we define $F : P \times [0,1] \to \Lambda(Q)$ by

$$F(x,t) := \mathrm{con}(\overline{tF_0(x)}, tF_1(x), \overline{tF_1(x)}, \dots, tF_{j-1}(x), \overline{tF_{j-1}(x)}, tF_j(x)).$$

The above concatenation is well-defined, since $F_0(x)(0) = \cdots = F_j(x)(0)$. For any $x \in P$, $\mathcal{E}(F_0(x)), \dots, \mathcal{E}(F_j(x)) < a/(2j)^2$. Thus, $\mathcal{E}(F(x,t)) < a$. Therefore, $F(P \times [0,1]) \subset \Lambda^a(Q)$. For any $x \in P$ and $k = 0, \dots, j$, there holds $0 \cdot F_k(x) = p_{f(x)}$, and therefore, $F(x,0) = p_{f(x)}$. This shows that F satisfies F-(i).

We check that F satisfies F-(ii). There holds $F(x,t) \in \Lambda_\partial(Q)$ for any $(x,t) \in P' \times [0,1]$, since $F(x,t)(1/2j) = F_0(x)(0) = f(x) \in \partial Q$. Hence, it is sufficient to show that $F(x,1) \in \Lambda_\partial(Q)$ for any $x \in P$. By Lemma 5.6 (iii), there exists a simplex σ of P such that $w_\sigma(x) = 1$. Let $k := \dim\sigma$. Then, $F(x,1)(k/j) = F_k(x)(1) \in \partial Q$. Hence, $F(x,1) \in \Lambda_\partial(Q)$. This completes the proof of $c^\Lambda(Q : \alpha) \leq 2jr(Q)$.

The proof of $c^\Omega(Q : \alpha) \leq 2(j+1)r(Q)$ is similar. Let us take $P' \subset P$ and $f : (P, P') \to (Q, \partial Q)$ so that $\alpha \in f_*(H_j(P, P'))$. It is sufficient to show that, if $a/(2j+2)^2 > r(Q)^2/2$, there exists a continuous map $F' : P \times [0,1] \to \Omega^a(Q)$ such that

F′-(i): For any $x \in P$, $F'(x,0) = l_{f(x)}$.

F′-(ii): For any $(x,t) \in P'' = P' \times [0,1] \cup P \times \{1\}$, $F'(x,t) \in \Omega_\partial(Q)$.

For each $k = 0, \dots, j$, we define $F'_k : P \to \Lambda^{a/(2j+2)^2}(Q)$ as in the proof of $c^\Lambda(Q : \alpha) \leq 2jr(Q)$. Then, we define F' by

$$F'(x,t) := \mathrm{con}(tF'_0(x), \overline{tF'_0(x)}, \dots, tF'_j(x), \overline{tF'_j(x)}).$$

Since $F'(x,t)(0) = f(x) = F'(x,t)(1)$, one can consider $F'(x,t)$ as an element in $\Omega(Q)$. It is easy to verify that $\mathcal{E}(F'(x,t)) < a$ for any $(x,t) \in P \times [0,1]$. Therefore, $F'(P \times [0,1]) \subset \Omega^a(Q)$. It is also easy to verify that F' satisfies F'-(i), (ii), in a similar manner as in the proof of $c^\Lambda(Q : \alpha) \leq 2jr(Q)$. □

5.3. Proof of Corollary 1.7. We conclude this section with a proof of Corollary 1.7.

Proof. The case $j = 1$ is easy, and therefore, omitted (see [14], pp.501–502). Hence, we may assume that M is simply connected. By the Hurewicz theorem, it is sufficient to show that if $H_j(M) \neq 0$, then there exists a nontrivial geodesic loop at p of length $\leq 2j\,\mathrm{diam}(M)$.

Let $\rho(M)$ be the injectivity radius of M. For any $\varepsilon < \rho(M)$, let $Q_\varepsilon := \{x \in M \mid \mathrm{dist}(x,p) \geq \varepsilon\}$. Then, it is clear that $r(Q_\varepsilon) \leq \mathrm{diam}(M) - \varepsilon < \mathrm{diam}(M)$. Moreover, $H_j(Q_\varepsilon, \partial Q_\varepsilon) \cong H_j(M) \neq 0$.

We apply Theorem 1.4 for Q_ε. Then, there exists a brake billiard trajectory γ_ε on Q_ε such that $\mathrm{length}(\gamma_\varepsilon) \le 2jr(Q_\varepsilon) < 2j\,\mathrm{diam}(M)$. We set $\tau_\varepsilon := \min\{t > 0 \mid \gamma_\varepsilon(t) \in \partial Q_\varepsilon\}$, and define $\Gamma_\varepsilon : [0,1] \to Q_\varepsilon$ by $\Gamma_\varepsilon(t) := \gamma_\varepsilon(\tau_\varepsilon t)$. Since $\Gamma_\varepsilon(0), \Gamma_\varepsilon(1) \in \partial Q_\varepsilon$, and $\mathrm{length}(\Gamma_\varepsilon) \le \mathrm{length}(\gamma_\varepsilon) < 2j\,\mathrm{diam}(M)$, a certain subsequence of $(\Gamma_\varepsilon)_\varepsilon$ converges to a geodesic loop $\Gamma : [0,1] \to M$ at p such that $\mathrm{length}(\Gamma) \le 2j\,\mathrm{diam}(M)$.

We have to check that Γ is nonconstant. Since $\dot{\Gamma}^+_\varepsilon(0)$ is perpendicular to ∂Q_ε and nonzero, $\Gamma_\varepsilon([0,1])$ intersects $S := \{x \in M \mid \mathrm{dist}(x,p) = \rho(M)\}$. Hence $\Gamma([0,1])$ also intersects S. Since $p \notin S$, Γ is nonconstant. □

6. Shortest periodic billiard trajectory in a convex body

In this section, we prove Theorem 1.8 and Theorem 1.9 using our method. A recent paper [3] obtained similar proofs based on the results in [8]. Several results in this section were already obtained in [8] in a more general setting. We include proofs of these results for the sake of completeness, although some arguments overlap with the arguments in [8].

First, let us introduce some notations. Let $K \subset \mathbb{R}^n$ be a convex body with C^∞ boundary.

- We abbreviate $c^\Omega(K : [K, \partial K])$ as $c^\Omega(K)$.
- $\mathcal{P}(K)$ denotes the set of periodic billiard trajectories in K.
- $\mathcal{P}^+(K)$ denotes the set consisting of piecewise geodesic curves $\gamma : S^1 \to \mathbb{R}^n$ such that $\gamma(S^1) + x \not\subset \mathrm{int} K$ for any $x \in \mathbb{R}^n$.
- For any $\nu \in \mathbb{R}^n$ and a compact set $S \subset \mathbb{R}^n$, $h(S : \nu) := \max\{s \cdot \nu \mid s \in S\}$.
- For any $q \in \partial K$, $\nu(q)$ denotes the unit vector that is outer normal to ∂K at q.

Lemma 6.1. *Let K be a convex body with C^∞ boundary, and $\gamma : S^1 \to \mathbb{R}^n$ be a piecewise geodesic curve. If there exists $\mathcal{N} \subset \mathbb{R}^n \setminus \{(0,\dots,0)\}$ such that $(0,\dots,0) \in \mathrm{conv}(\mathcal{N})$ and $h(K : \nu) \le h(\gamma(S^1) : \nu)$ for any $\nu \in \mathcal{N}$, then $\gamma \in \mathcal{P}^+(K)$.*

Proof. Take $x \in \mathbb{R}^n$ arbitrarily. Since $(0,\dots,0) \in \mathrm{conv}(\mathcal{N})$, there exists $\nu \in \mathcal{N}$ such that $x \cdot \nu \ge 0$. Thus, $h(\gamma(S^1) + x : \nu) \ge h(\gamma(S^1) : \nu) \ge h(K : \nu)$. Since $\nu \ne 0$, this shows that $\gamma(S^1) + x \not\subset \mathrm{int} K$. □

Lemma 6.2. *Any $\gamma \in \mathcal{P}(K)$ satisfies the assumption in Lemma 6.1 with $\mathcal{N} := \{\nu(\gamma(t)) \mid t \in \mathcal{B}_\gamma\}$. In particular, $\mathcal{P}(K) \subset \mathcal{P}^+(K)$.*

Proof. For any $t \in \mathcal{B}_\gamma$, there holds $h(K : \nu(\gamma(t))) = \gamma(t) \cdot \nu(\gamma(t))$ since K is convex. Hence $h(K : \nu) = h(\gamma(S^1) : \nu)$ for any $\nu \in \mathcal{N}$.

Suppose that $(0,\dots,0) \notin \mathrm{conv}(\mathcal{N})$. Since $\mathcal{N}$ is a finite set, there exists $x \in \mathbb{R}^n$ such that $x \cdot \nu > 0$ for any $\nu \in \mathcal{N}$. Since $\ddot{\gamma} \equiv 0$ on $S^1 \setminus \mathcal{B}_\gamma$, there exists $t \in \mathcal{B}_\gamma$

such that $x \cdot (\dot{\gamma}^-(t) - \dot{\gamma}^+(t)) \leq 0$. On the other hand, it is easy to see that $\nu(\gamma(t)) = \dot{\gamma}^-(t) - \dot{\gamma}^+(t)/|\dot{\gamma}^-(t) - \dot{\gamma}^+(t)|$. Thus, we have $x \cdot \nu(\gamma(t)) \leq 0$. This is a contradiction, thus $(0, \dots, 0) \in \mathrm{conv}(\mathcal{N})$. □

The following proposition is a key step in the proof.

Proposition 6.3. *For any $\gamma \in \mathcal{P}^+(K)$, there holds $c^\Omega(K) \leq \mathrm{length}(\gamma)$.*

Proof. It is sufficient to show that, for any $a > \mathrm{length}(\gamma)^2/2$ and $\delta > 0$, the homomorphism

$$H_n(\mathrm{int}K, \mathrm{int}K \setminus K(\delta)) \to H_n(\Omega^a(\mathrm{int}K), \Omega^a(\mathrm{int}K) \cap \Omega_\delta(\mathrm{int}K))$$

is zero. By the excision property, this is equivalent to show that

$$H_n(\mathbb{R}^n, \mathbb{R}^n \setminus K(\delta)) \to H_n(\Omega^a(\mathbb{R}^n), \Omega^a(\mathbb{R}^n) \setminus \Omega(K(\delta)))$$

is zero. By changing parameters of γ if necessary, we may assume that $\mathcal{E}(\gamma) = \mathrm{length}(\gamma)^2/2$. Therefore, $\mathcal{E}(\gamma) < a$.

Let us set $B_R := \{x \in \mathbb{R}^n \mid |x| \leq R\}$ for any $R > 0$. We define $F : B_R \times [0,1] \to \Omega^a(\mathbb{R}^n)$ by $F(w,s)(t) := w + s\gamma(t)$. When R is sufficiently large, $w + s\gamma(S^1) \not\subset K(\delta)$ for any $w \in \partial B_R$ and $0 \leq s \leq 1$. Moreover, $w + \gamma(S^1) \not\subset K(\delta)$ for any $w \in B_R$, since $\gamma \in \mathcal{P}^+(K)$. Thus, setting $P := B_R \times [0,1]$ and $P' := \partial B_R \times [0,1] \cup B_R \times \{1\}$, we have

$$F : (P, P') \to (\Omega^a(\mathbb{R}^n), \Omega^a(\mathbb{R}^n) \setminus \Omega(K(\delta))).$$

Setting $i : (B_R, \partial B_R) \to (P, P'); x \mapsto (x, 0)$, we have the commutative diagram

$$\begin{array}{ccc} H_n(B_R, \partial B_R) & \xrightarrow{i_*} & H_n(P, P') \\ \downarrow & & \downarrow F_* \\ H_n(\mathbb{R}^n, \mathbb{R}^n \setminus K(\delta)) & \longrightarrow & H_n(\Omega^a(\mathbb{R}^n), \Omega^a(\mathbb{R}^n) \setminus \Omega(K(\delta))). \end{array}$$

Since $K(\delta)$ is also convex, the left vertical arrow is an isomorphism. On the other hand, $i_* = 0$. Thus, the bottom homomorphism is zero. □

Corollary 6.4. *Let us define $\mu_P^+(K) := \inf\{\mathrm{length}(\gamma) \mid \gamma \in \mathcal{P}^+(K)\}$. Then, $c^\Omega(K) = \mu_P(K) = \mu_P^+(K)$.*

Proof. Lemma 5.4 shows $c^\Omega(K) \geq \mu_P(K)$. $\mathcal{P}(K) \subset \mathcal{P}^+(K)$ shows $\mu_P(K) \geq \mu_P^+(K)$. Proposition 6.3 shows $\mu_P^+(K) \geq c^\Omega(K)$. □

Remark 6.5. The identity $c^\Omega(K) = \mu_P(K)$ implies that there exists a shortest periodic billiard trajectory in K, since Lemma 5.4 shows that there exists a periodic billiard trajectory γ in K such that $\mathrm{length}(\gamma) = c^\Omega(K)$.

The identity $\mu_P(K) = \mu_P^+(K)$ can be considered as a variational characterization of μ_P. The same result is established in [8] (see also [3], Theorem 2.1). As an immediate consequence, we can recover the following result, which was already obtained in Proposition 1.4 [5] (see also [3] Section 2.2).

Corollary 6.6 ([5]). *Let $K_1 \subset K_2$ be convex bodies with C^∞ boundaries. Then, $\mu_P(K_1) \le \mu_P(K_2)$.*

Proof. It is obvious that $\mathcal{P}^+(K_2) \subset \mathcal{P}^+(K_1)$. Then, we have

$$\mu_P(K_1) = \mu_P^+(K_1) \le \mu_P^+(K_2) = \mu_P(K_2).$$

□

We also need Lemma 6.7 to determine when equality holds in Theorem 1.8 and Theorem 1.9.

Lemma 6.7. *Suppose that $\gamma \in \mathcal{P}^+(K)$ satisfies* $\text{length}(\gamma) = \mu_P^+(K)$*, and $|\dot\gamma(t)|$ is constant for all t such that $\dot\gamma(t)$ exists. Then, up to parallel displacement, $\gamma \in \mathcal{P}(K)$.*

Proof. For any $\varepsilon > 0$, we set $\gamma_\varepsilon(t) := (1-\varepsilon)\gamma(t)$. Since $\text{length}(\gamma_\varepsilon) < \text{length}(\gamma) = \mu_P^+(K)$, there holds $\gamma_\varepsilon \notin \mathcal{P}^+(K)$. There exists $x_\varepsilon \in \mathbb{R}^n$ such that $x_\varepsilon + \gamma_\varepsilon(S^1) \subset \text{int}K$ for any $\varepsilon > 0$, thus by parallel displacement, we may assume that $\gamma(S^1) \subset K$. We show that $\gamma \in \mathcal{P}(K)$.

Take $0 = t_0 < t_1 < \cdots < t_m = 1$ so that $\gamma|_{[t_{j-1},t_j]}$ are geodesics and $\dot\gamma^-(t_j) \ne \dot\gamma^+(t_j)$ for all $1 \le j \le m$. We set $J := \{1 \le j \le m \mid \gamma(t_j) \in \partial K\}$. For each $j \in J$, let us abbreviate $\nu(\gamma(t_j))$ as ν_j. By convexity of K, $h(K : \nu_j) = \gamma(t_j) \cdot \nu_j$ for each $j \in J$.

Let $\mathcal{N} := \{\nu_j \mid j \in J\}$. If $(0,\dots,0) \notin \text{conv}(\mathcal{N})$, there exists $x \in \mathbb{R}^n$ such that $x \cdot \nu_j < 0$ for any $j \in J$. Thus, $\gamma(S^1) + cx \subset \text{int}K$ for sufficiently small $c > 0$. This is impossible since $\gamma \in \mathcal{P}^+(K)$. Thus, we have shown $(0,\dots,0) \in \text{conv}(\mathcal{N})$.

We show that $J = \{1,\dots,m\}$. If $J \subsetneq \{1,\dots,m\}$, there exists $\gamma' : S^1 \to K$ such that $\text{length}(\gamma') < \text{length}(\gamma)$ and $\gamma'(S^1) \supset \{\gamma(t_j) \mid j \in J\}$. For each $j \in J$, one has

$$h(K : \nu_j) = \gamma(t_j) \cdot \nu_j \le h(\gamma'(S^1) : \nu_j).$$

Then, Lemma 6.1 implies $\gamma' \in \mathcal{P}^+(K)$. This is impossible since γ has the shortest length in $\mathcal{P}^+(K)$.

To prove $\gamma \in \mathcal{P}(K)$, it is sufficient to check that γ satisfies the law of reflection at every t_j. If this is not the case, i.e., $\dot\gamma^+(t_j) - \dot\gamma^-(t_j)$ is not a multiple of ν_j for some j, there exists $v \in T_{\gamma(t_j)}\partial K$ such that

$$|\gamma(t_j) - \gamma(t_{j-1})| + |\gamma(t_{j+1}) - \gamma(t_j)| > |\gamma(t_j) + v - \gamma(t_{j-1})| + |\gamma(t_{j+1}) - \gamma(t_j) - v|.$$

Define $\gamma' : S^1 \to \mathbb{R}^n$ so that

$$\gamma'(t_i) = \begin{cases} \gamma(t_j) + v & (i = j) \\ \gamma(t_i) & (i \neq j) \end{cases}$$

and $\gamma'|_{[t_{i-1}, t_i]}$ are geodesics for all $1 \le i \le m$. Then, $\mathrm{length}(\gamma') < \mathrm{length}(\gamma)$. It is easy to check $h(\gamma'(S^1) : \nu_i) \ge h(K : \nu_i)$ for any $1 \le i \le m$, and thus, Lemma 6.1 implies $\gamma' \in \mathcal{P}^+(K)$. This is impossible since γ has the shortest length in $\mathcal{P}^+(K)$. □

For any two curves $\gamma_i : S^1 \to \mathbb{R}^n$ $(i = 1, 2)$, we define $\gamma_1 + \gamma_2 : S^1 \to \mathbb{R}^n$ by $\gamma_1 + \gamma_2(t) := \gamma_1(t) + \gamma_2(t)$. The following lemma would be clear from the definition of $\mathcal{P}^+$.

Lemma 6.8. *If $\gamma_i(S^1) \notin \mathcal{P}^+(K_i)$ for $i = 1, 2$, one has $\gamma_1 + \gamma_2 \notin \mathcal{P}^+(K_1 + K_2)$.*

Now, we can prove Theorem 1.8.

Proof of Theorem 1.8. Let $a_j := \frac{\mu_P(K_j)}{\mu_P(K_1) + \mu_P(K_2)}$. If $\mathrm{length}(\gamma) < \mu_P(K_1) + \mu_P(K_2)$, we have the following inequality for each $j = 1, 2$.

$$\mathrm{length}(a_j\gamma) = a_j \cdot \mathrm{length}(\gamma) < \mu_P(K_j) = \mu_P^+(K_j).$$

Then, $a_j\gamma \notin \mathcal{P}^+(K_j)$. By Lemma 6.8, $\gamma = a_1\gamma + a_2\gamma \notin \mathcal{P}^+(K_1 + K_2)$. Thus, we have shown that $\mu_P^+(K_1 + K_2) \ge \mu_P(K_1) + \mu_P(K_2)$. By Corollary 6.4, we get

$$\mu_P(K_1 + K_2) = \mu_P^+(K_1 + K_2) \ge \mu_P(K_1) + \mu_P(K_2). \tag{6.1}$$

We have to show that the following two conditions are equivalent.

(i): $\mu_P(K_1 + K_2) = \mu_P(K_1) + \mu_P(K_2)$.

(ii): There exists a closed curve γ which, up to parallel displacement and scaling, is the shortest periodic billiard trajectory in both K_1 and K_2.

(i) $\Longrightarrow$ (ii): There exists $\gamma \in \mathcal{P}(K_1 + K_2)$ such that $\mathrm{length}(\gamma) = \mu_P(K_1 + K_2)$. If $a_1\gamma \notin \mathcal{P}^+(K_1)$, one has $(a_1 + \varepsilon)\gamma \notin \mathcal{P}^+(K_1)$ for sufficiently small $\varepsilon > 0$. On the other hand, $(a_2 - \varepsilon)\gamma \notin \mathcal{P}^+(K_2)$ since $(a_2 - \varepsilon)\mathrm{length}(\gamma) < \mu_P(K_2)$. Thus, $\gamma \notin \mathcal{P}^+(K_1 + K_2)$, which is a contradiction. Therefore, $a_1\gamma \in \mathcal{P}^+(K_1)$. Since $\mathrm{length}(a_1\gamma) = \mu_P(K_1)$, Lemma 6.7 implies $a_1\gamma \in \mathcal{P}(K_1)$ up to parallel displacement. We can prove $a_2\gamma \in \mathcal{P}(K_2)$ in the same manner, and thus, (ii) holds.

(ii) $\Longrightarrow$ (i): Take $\gamma : S^1 \to \mathbb{R}^n$ as in (ii). For $j = 1, 2$, let γ_j be a shortest periodic billiard trajectory on K_j, which is obtained by parallel displacement and scaling of γ. We may assume that $\gamma = \gamma_1 + \gamma_2$. Then, $\mathrm{length}(\gamma) = \mathrm{length}(\gamma_1) + \mathrm{length}(\gamma_2) = \mu_P(K_1) + \mu_P(K_2)$.

It is easy to see that $\mathcal{B}_{\gamma_1} = \mathcal{B}_{\gamma_2}$. Let us denote it as $\mathcal{B}$. For each $t \in \mathcal{B}$, $\nu(t) := \dot\gamma^-(t) - \dot\gamma^+(t)/|\dot\gamma^-(t) - \dot\gamma^+(t)|$ is outer normal to ∂K_j at $\gamma_j(t)$ for $j = 1, 2$.

Let us set $\mathcal{N} := \{\nu(t) \mid t \in \mathcal{B}\}$. By Lemma 6.2, we have $(0,\dots,0) \in \operatorname{conv}(\mathcal{N})$ and $h(K_j : \nu) = h(\gamma_j(S^1) : \nu)$ for any $\nu \in \mathcal{N}$, $j = 1, 2$. Then, for any $\nu \in \mathcal{N}$

$$h(K_1 + K_2 : \nu) = h(K_1 : \nu) + h(K_2 : \nu) = h(\gamma_1(S^1) : \nu) + h(\gamma_2(S^1) : \nu) \\ = h(\gamma(S^1) : \nu).$$

By Lemma 6.1, $\gamma \in \mathcal{P}^+(K_1 + K_2)$. Hence,

$$\mu_P(K_1 + K_2) = \mu_P^+(K_1 + K_2) \le \operatorname{length}(\gamma) = \mu_P(K_1) + \mu_P(K_2).$$

Combined with (6.1), (i) is proved. □

To prove Theorem 1.9, we need the following lemma.

Lemma 6.9. *Let B be a ball in $\mathbb{R}^n$ with radius $r > 0$. Then, any $\gamma \in \mathcal{P}(B)$ satisfies* $\operatorname{length}(\gamma) \ge 4r$*, and equality holds if and only if γ is a bouncing ball orbit. In particular, $\mu_P(B) = 4r$.*

Proof. Let $k := \sharp\mathcal{B}_\gamma$. Then, one has $\operatorname{length}(\gamma) = 2kr\sin(\pi j/k)$ for some $1 \le j \le k-1$. Then, the lemma follows from short computations. □

Proof of Theorem 1.9. Let K be a convex body, and B be the largest ball contained in K. Since the radius of B is $r(K)$, Corollary 6.6 and Lemma 6.9 imply $\mu_P(K) \ge \mu_P(B) = 4r(K)$.

Suppose that $\mu_P(K) = 4r(K)$, and let γ be the shortest periodic billiard trajectory in K. Then, $\gamma \in \mathcal{P}(K) \subset \mathcal{P}^+(K) \subset \mathcal{P}^+(B)$, and $\operatorname{length}(\gamma) = 4r(K) = \mu_P(B)$. Then, Lemma 6.7 shows that $\gamma \in \mathcal{P}(B)$ up to parallel displacement. By Lemma 6.9, γ is a bouncing ball orbit. In particualr, γ is orthogonal to ∂K at bouncing points. Thus, K is contained in a slab of thickness $\operatorname{length}(\gamma)/2 = 2r(K)$. Hence $\operatorname{width}(K) = 2r(K)$.

Suppose that $\operatorname{width}(K) = 2r(K)$. Then, K is contained in a slab S of thickness $2r(K)$. Let γ be a bouncing ball orbit on S, i.e., γ is the shortest orbit that touches both connected components of ∂S. Then, it is easy to see that $\gamma \in \mathcal{P}^+(S) \subset \mathcal{P}^+(K)$. Thus, $\mu_P(K) = \mu_P^+(K) \le \operatorname{length}(\gamma) = 4r(K)$. □

Acknowledgements. The author would like to thank the referees for their useful comments, in particular, Remark 2.13. The author is supported by JSPS KAKENHI Grant No. 25800041.

References

[1] A. Abbondandolo, P. Majer, *Lectures on the Morse complex for infinite-dimensional manifolds*, Morse theoretic methods in nonlinear analysis and in symplectic topology, 1–74, Springer, Dordrecht, 2006. Zbl 1089.37007 MR 2276948

[2] A. Abbondandolo, M. Schwarz, *A smooth pseudo-gradient for the Lagrangian action functional*, Adv. Nonlinear Stud. **9** (2009), no. 4, 597–623. Zbl 1185.37145 MR 2560122

[3] A. Akopyan, A. Balitskiy, R. Karasev, A. Sharipova, *Elementary approach to closed billiard trajectories in asymmetric normed spaces*, arXiv:1401.0442v3, 2014.

[4] S. Artstein-Avidan, Y. Ostrover, *A Brunn-Minkowski inequality for symplectic capacities of convex domains*, Int. Math. Res. Not. IMRN 2008, no. 13, Art. ID rnn044, 31 pp. Zbl 1149.52006 MR 2436562

[5] S. Artstein-Avidan, Y. Ostrover, *Bounds for Minkowski billiard trajectories in convex bodies*, Int. Math. Res. Not. IMRN 2014, no. 1, 165–193. Zbl 06340328 MR 3158530

[6] V. Benci, *Periodic solutions of Lagrangian systems on a compact manifold*, J. Differential Equations. **63** (1986), no. 2, 135–161. Zbl 0605.58034 MR 848265

[7] V. Benci, F. Giannoni, *Periodic bounce trajectories with a low number of bounce points*, Ann. Inst. H. Poincaré Anal. Non Linéaire. **6** (1989), no. 1, 73–93. Zbl 0667.34054 MR 984148

[8] D. Bezdek, K. Bezdek, *Shortest billiard trajectories*, Geom. Dedicata. **141** (2009), 197–206. Zbl 1169.52002 MR 2520072

[9] A. Floer, H. Hofer, K. Wysocki, *Applications of symplectic homology I*, Math. Z. **217** (1994), no. 4, 577–606. Zbl 0869.58012 MR 1306027

[10] M. Ghomi, *Shortest periodic billiard trajectories in convex bodies*, Geom. Funct. Anal. **14** (2004), no. 2, 295–302. Zbl 1059.52006 MR 2060196

[11] H. Hofer, E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser, Basel, 1994. Zbl 0805.58003 MR 1306732

[12] K. Irie, *Symplectic homology of disc cotangent bundles of domains in Euclidean space*, J. Symplectic Geom. **12** (2014), no. 3, 511–552. Zbl 1302.53095 MR 3248667

[13] W. Klingenberg, *Lectures on closed geodesics*, Grundlehren der Mathematischen Wissenschaften vol. 230, Springer-Verlag, Berlin-New York, 1978. Zbl 0397.58018 MR 478069

[14] R. Rotman, *The length of a shortest geodesic loop at a point*, J. Differential Geom. **78** (2008), no. 3, 497–519. Zbl 1143.53038 MR 2396252

[15] M. Schwarz, *On the action spectrum for closed symplectically aspherical manifolds*, Pacific J. Math. **193** (2000), no. 2, 419–461. Zbl 1023.57020 MR 1755825

[16] J. Weber, *Perturbed closed geodesics are periodic orbits: index and transversality*, Math. Z. **241** (2002), no. 1, 45–82. Zbl 1037.53060 MR 1930985

Received June 29, 2013

K. Irie, Research Institute for Mathematical Sciences, Kyoto University,
Kyoto 606-8502, Japan

E-mail: iriek@kurims.kyoto-u.ac.jp