# A QUADRATIC LOWER BOUND FOR THE NUMBER OF MINIMAL GEODESICS

BERND AMMANN AND CLARA LÖH

ABSTRACT. A minimal geodesic on a Riemannian manifold is a geodesic defined on $\mathbb{R}$ that lifts to a globally distance minimizing curve on the universal covering. Bangert proved that there is a lower bound for the number of geometrically distinct minimal geodesics of closed Riemannian manifolds that is linear in the first Betti number, using the stable norm unit ball on the first homology. We refine this method to obtain a quadratic lower bound. For example, on the 3-dimensional torus with an arbitrary Riemannian metric we improve the lower bound from 3 to 15. We distinguish between different types of minimal geodesics and we show that our lower estimate for the number of homologically non-homoclinic minimal geodesics is sharp.

## 1. INTRODUCTION

1.1. **Guiding question of the article.** We study the following classical geometric counting problem (see Section 2 for the precise definitions):

**Question 1.1.** *How many geometrically distinct minimal geodesics does a closed connected Riemannian manifold need to have?*

It is well-known that a minimal geodesic exists if and only if $\pi_1(M)$ is infinite (Proposition A.1). Bangert proved that a closed connected Riemannian manifold has at least $\dim_{\mathbb{R}} H_1(M;\mathbb{R})$ minimal geodesics [6]. For the circle, there clearly is exactly one. For 2-dimensional tori [19] (see Example 9.1) and closed connected surfaces of genus at least 2 [28, 21] (see Example 9.2), there are always infinitely many. Similarly, if $M$ is a closed connected Riemannian manifold with non-elementary Gromov-hyperbolic fundamental group, then on the universal covering $\widetilde{M}$ each two different points of the boundary of $\widetilde{M}$ at infinity can be joined by a minimal geodesic, and thus there are uncountably many minimal geodesics on $M$ [17, Section 7.5].

In contrast, in dimension at least 3, examples with very few minimal geodesics have been constructed by Hedlund [19, Section 9] and others [6, Section 5][1]. Hedlund's work is sometimes misunderstood and it is erroneously claimed that he constructed metrics with only three geometrically distinct minimal geodesics on the 3-dimensional torus. Although it is easy to see that Hedlund's construction does not have this property, it was not clear whether one might find other Riemannian metrics on the 3-torus that have only three minimal geodesics.

We answer this question negatively: We show the existence of at least 15 geometrically distinct minimal geodesics on the 3-dimensional torus with an arbitrary Riemannian metric, see Corollary 1.5.

*Date*: April 11, 2024. © B. Ammann, C. Löh, 2023. This work was supported by the CRC 1085 *Higher Invariants* (Universität Regensburg), funded by the DFG. B. A. was also supported by the SPP 2026 (Geometry at infinity), funded by the DFG.

1.2. **A quadratic lower bound.** To state our main result, we introduce the following constants:

**Definition 1.2.** Let $b \in \mathbb{N}$ and let $\mathrm{CS}(b)$ be the set of all centrally symmetric polytopes in $\mathbb{R}^b$ with non-empty interior. If $P \in \mathrm{CS}(b)$, we write $V(P)$ and $E(P)$ for the number of vertices and edges of $P$, respectively. We set

$$\begin{aligned} \mathrm{E}_{\min}(b) &:= \min_{P \in \mathrm{CS}(b)} E(P), \\ \mathrm{VE}_{\min}(b) &:= \min_{P \in \mathrm{CS}(b)} \left( \frac{V(P)}{2} + E(P) \right). \end{aligned}$$

Note that the minimal number of vertices is $2b$, and thus $\mathrm{VE}_{\min}(b) \geq b + \mathrm{E}_{\min}(b)$. We do not know whether there are $b \in \mathbb{N}$ with $\mathrm{VE}_{\min}(b) > b + \mathrm{E}_{\min}(b)$, see Remark 1.4 (a) below.

**Theorem 1.3.** *Let $M$ be a closed connected Riemannian manifold and let $b := \dim_{\mathbb{R}} H_1(M;\mathbb{R})$. Then $M$ admits at least $\mathrm{VE}_{\min}(b)$ geometrically distinct minimal geodesics. Moreover,*

$$\mathrm{VE}_{\min}(b) \geq \begin{cases} 2 \cdot b^2 - b & \text{if } b \leq 3 \\ b^2 + 2 \cdot b + 1 & \text{if } b \geq 4. \end{cases}$$

We briefly outline the argument (which is presented in full detail in Section 6, Section 7, and Subsection 8.1). See also Subsections 3.4 and 3.5 for precise definitions of exposed edges and exposed points. The key tools in the proof are the (closed) unit ball $B \subset H_1(M;\mathbb{R})$ of the stable norm on $H_1(M;\mathbb{R})$ and the construction of minimizing geodesics on the universal covering $\widetilde{M}$ of $M$ from sequences of finite-length minimizing geodesics.

Bangert proved that every antipodal pair of exposed points of $B$ leads to a minimal geodesic whose asymptotic direction is controlled by the underlying exposed point [6, Theorem 4.4]. If $B$ has infinitely many exposed points, then this already shows that $M$ has infinitely many geometrically distinct minimal geodesics.

If $B$ has only finitely many exposed points, then $B$ is a compact convex centrally symmetric polytope in $H_1(M;\mathbb{R})$, and the exposed points are the vertices of this polytope. We refine Bangert's method to construct additional minimal geodesics, whose asymptotic behaviour is controlled by the exposed edges of $B$. Moreover, we show that these are indeed geometrically distinct. To complete the proof, we give a lower bound for the number of vertices and edges in centrally symmetric compact convex polytopes, which gives the claimed estimate for $\mathrm{VE}_{\min}$.

**Remark 1.4.**

(a) We have $\mathrm{VE}_{\min}(b) \geq b + \mathrm{E}_{\min}(b)$ (Corollary 7.2). For $\mathrm{E}_{\min}$, we conjecture the lower bound $\mathrm{E}_{\min}(b) \geq 2 \cdot b \cdot (b-1)$ for all $b$. The conjecture is equivalent to saying that the minimum in $\mathrm{E}_{\min}$ is attained by the cross-polytope. The *cross-polytope* is, by definition, the unit norm ball of the $\ell^1$-norm. The conjecture is equivalent to $\mathrm{VE}_{\min}(b) = b + \mathrm{E}_{\min}(b)$ (Corollary 7.2).

(b) In a forthcoming work [3], we will show that $\mathrm{E}_{\min}(b) \geq b \cdot (b+2)$ for even $b \geq 4$ and that this is sharp for $b = 4$. For odd $b \geq 5$ will show that $\mathrm{E}_{\min}(b) \geq b \cdot (b+2) - 1$.

**Corollary 1.5.** *Let $n \in \mathbb{N}_{\geq 3}$ and let $g$ be a Riemannian metric on the $n$-torus $T^n$. Then $(T^n, g)$ admits at least*

$$\begin{cases} 15 & \text{if } n = 3 \\ (n+1)^2 & \text{if } n \geq 4 \end{cases}$$

*geometrically distinct minimal geodesics.*

*Proof.* We apply Theorem 1.3 and use that $\dim_{\mathbb{R}} H_1(T^n; \mathbb{R}) = n$. □

Thus, each Riemannian 3-torus admits at least 15 geometrically distinct minimal geodesics. In particular, this applies to the Hedlund examples [19, Section 9][6, Section 5] on $T^3$ and shows that the corresponding exercise in the book by D. Burago, Y. Burago, and S. Ivanov [12, Exercise 8.5.16] is not solvable. The latter is already evident from Bangert's study of Hedlund metrics [6, Section 5]. Our result implies that there is also no other Riemannian metric on the 3-torus that admits only three geometrically distinct minimal geodesics.

Similarly, in combination with Remark 1.4 (b), we get for the 4-dimensional torus that besides the four minimal geodescis detected by Bangert [6, Theorem 4.4] there are at least 24 minimal geodesics of of another type (namely "homologically non-homoclinic and $\mathbb{R}$-homologically minimal"); in particular, the latter minimal geodesics are geometrically distinct from the ones detected by Bangert. We will also discuss that our bound is sharp as a lower bound on the number of homologically non-homoclinic and $\mathbb{R}$-homologically minimal geodesics.

1.3. **A refinement.** Theorem 1.3 can be strengthened, by distinguishing between different types of minimal geodesics. This leads to the following refinement, stated in Theorem 1.6. The terms "homologically homoclinic", "homologically heteroclinic" and "homologically non-homoclinic" will be defined in Definition 5.13; "$\mathbb{R}$-homologically minimal" geodesics are introduced in Section 2.5. The minimal geodesics obtained by Bangert are homologically homoclinic and $\mathbb{R}$-homologically minimal.

**Theorem 1.6** (refinement of Theorem 1.3)**.** *Let $M$ be a closed connected Riemannian manifold and let $b := \dim_{\mathbb{R}} H_1(M; \mathbb{R})$. Then, at least one of the following statements holds:*

(a) *There are infinitely many geometrically distinct minimal geodesics that are homologically homoclinic and $\mathbb{R}$-homologically minimal.*
(b) *There are at least $\mathrm{E}_{\min}(b)$ geometrically distinct minimal geodesics that are homologically non-homoclinic and $\mathbb{R}$-homologically minimal, and there are at least $b$ geometrically distinct geodesics that are homologically homoclinic and $\mathbb{R}$-homologically minimal.*

*If the stable norm unit ball is not a polytope, then (a) is satisfied. If the stable norm unit ball is a polytope $B$, then one can replace $\mathrm{E}_{\min}(b)$ by $E(B)$ and the number of homologically homoclinic minimal geodesics is at least $1/2 \cdot V(B)$.*

We prove this in a stronger form in Proposition 6.8 and Section 8.2.

1.4. **Discussion of sharpness.** Let us discuss the sharpness of the previous results. We assume in this section, that the norm ball for the stable norm is a polytope, as otherwise Bangert's results already provide an infinite number of geometrically distinct minimal geodesics. Moreover, we assume that we are not in the case (A) of Proposition 6.8 in which there are uncountably many geometrically distinct homologically homoclinic minimal geodesics.

Our construction of minimal geodesics makes substantial use of the fact that we are working with vertices and edges of the stable norm unit ball. We do not expect that higher-dimensional faces in the boundary lead to further minimal geodesics.

It is natural to ask whether the lower bounds in Theorems 1.3 and 1.6 are sharp, i.e., whether on a given closed connected manifold $M$ there exist Riemannian metrics $g$ whose number of minimal geodesics attains this lower bound.

- As discussed above, the bound $\mathrm{VE}_{\min}$ is far from being optimal if one considers minimal geodesics on manifolds with non-elementary Gromov-hyperbolic fundamental group.

The sharpness is better controlled if one only considers $\mathbb{R}$-homologically minimal geodesics and special conditions (see Section 9.2 for details on the Hedlund examples):

- In the special case $\dim M \geq 3$ and $b = 2$, we explain in Example 9.4 the construction of Hedlund metrics for which the lower bound in Theorem 1.6 (b) on homologically non-homoclinic and $\mathbb{R}$-homologically minimal geodesics is attained. In this case, $\mathrm{E}_{\min}(2) = 4$ and the Hedlund metrics defined with respect to two generators of $H_1(M;\mathbb{Z})_{\mathbb{R}}$ have precisely four non-homoclinic $\mathbb{R}$-homologically minimal geodesics.
- Similar constructions are possible in the case $\dim M \geq 3$ and $b = 2$, for Hedlund metrics with respect to $\gamma_1, \dots, \gamma_k \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ with $k \geq 3$ for $H_1(M;\mathbb{Z})_{\mathbb{R}}$. We assume that $\gamma_1, \dots, \gamma_k$ span $H_1(M, R)$ (as a vector space). Then, the convex hull $B$ of $\{\pm\gamma_j \mid j \in \{1, \dots, k\}\}$ is a convex centrally symmetric $2\ell$-gon, for some $\ell \in \{2, \dots, k\}$. An associated Hedlund metric will have $B$ as the stable norm unit ball. By removing some of the $\gamma_i$ we can achieve $k = \ell$. Then the number of homologically non-homoclinic $\mathbb{R}$-homologically minimal geodesics is precisely $E(B) = 2k$. Thus we obtain a sharp estimate for homologically non-homoclinic $\mathbb{R}$-homologically minimal geodesics in this situation as well; see Example 9.5 for details.
- If we strengthen the hypotheses of the two previous items to $\pi_1(M) \cong \mathbb{Z}^2$, then a geodesic is minimal if and only if it is $\mathbb{R}$-homologically minimal. Thus, in this case, all statements above also hold for "homlogically non-homoclinic minimal geodesics" instead of "homologically non-homoclinic $\mathbb{R}$-homologically minimal geodesics".
- Sharpness is problematic for $b \geq 3$ and in the homoclinic case as the Hedlund examples exhibit unwanted "side shift effects", e.g., minimal geodesics of the type of Example 9.4 (1) and (2), see also Subsection 9.4. This affects the type of geodesics detected by Bangert's method.
- Even in the case of the 3-dimensional torus $T^3$ the minimal number of minimal geodesics is unknown. The known types of Hedlund examples have infinitely many minimal geodesics due to "side shift effects", but we expect that more refined construction methods can reduce them to 15 homoclinic and 12 heteroclinic minimal geodesics. In contrast to this, Bangert's and our lower bounds show the existence of three homologically homoclinic and 12 homologically non-homoclinic minimal geodesics. As the heteroclinic geodesics in such Hedlund examples are homologically non-homoclinic, the lower bound for the number of homologically non-homoclinic minimal geodesics is thus optimal. See again Subsection 9.4 for related statements.

A related problem is to study the consequences of equality in our estimates. We recall that a finite number of minimal geodesics is only possible if the stable norm unit ball is a polytope (Remark 8.3).

**Theorem 1.7** (Section 8.3)**.** *If $M$ is a closed connected Riemannian manifold with $b = \dim_{\mathbb{R}} H_1(M;\mathbb{R})$ and precisely $b+\mathrm{E}_{\min}(b)$ minimal geodesics, then the stable norm unit ball of $M$ is a cross-polytope and $\mathrm{E}_{\min}(b) = 2\cdot b^2 - 2\cdot b$. Furthermore, then all minimal geodesics are $\mathbb{R}$-homologically minimal and homologically exposed; among them $b$ are homologically homoclinic and $\mathrm{E}_{\min}(b)$ are homologically heteroclinic.*

In particular, the first phrase of the theorem implies: If the conjectured lower bound $\mathrm{E}_{\min}(b) \geq 2\cdot b\cdot(b-1)$ does not hold, then there would be at least $b+\mathrm{E}_{\min}(b)+1$ minimal geodesics.

**Theorem 1.8** (Section 8.4)**.** *Let $M$ be a closed connected Riemannian manifold with $b = \dim_{\mathbb{R}} H_1(M;\mathbb{R})$ whose stable norm unit ball $B$ is a polytope. We assume that $M$ has exactly $V(B)/2 + E(B)$ geometrically distinct minimal geodesics. Then all of them are $\mathbb{R}$-homologically minimal and homologically exposed; among them, $V(B)/2$ are homologically homoclinic and $E(B)$ are homologically heteroclinic.*

1.5. **Open problems.** The following problem seems to be open:

**Open Question 1.9.** *Do there exist closed connected Riemannian manifolds $M$ that satisfy $\dim_{\mathbb{R}} H_1(M;\mathbb{R}) \geq 2$ and have only finitely many geometrically distinct minimal geodesics?*

As discussed above, we expect that Hedlund examples on the 3-torus can be improved in a way such that one can show that they have at most $12 + 15 = 27$ geometrically distinct minimal geodesics.

Informally speaking, the real strength of the Hedlund examples [19, Section 9][6, Section 5][1, 2] lies in the fact that they have "few" minimal geodesics in an asymptotic sense. Many of these examples have an infinite number of minimal geodesics, however with only finitely many asymptotic types.

In order to turn this into a precise statement, we define what it means that a geodesic $\tau : \mathbb{R} \to M$ is "asymptotic in the future/past direction" to the geodesic $\gamma : \mathbb{R} \to M$. For simplicity, we supose that both geodesics are parametrized by arclength. We define $T^{\pm}\gamma$ as

$$T^{\pm}\gamma \coloneqq \left\{\dot\gamma(t) \;\middle|\; t \in \mathbb{R}\right\} \cup \left\{-\dot\gamma(t) \;\middle|\; t \in \mathbb{R}\right\},$$

which is a subset of the unit tangent bundle $SM$. We choose a Riemannian metric on the unit tangent bundle $SM$, which defines a distance function $d^{SM}$ on $SM$. Finally, for $v \in SM$ we write $d^{SM}(v, T^{\pm}\gamma) \coloneqq \inf\{d^{SM}(v, w) \mid w \in T^{\pm}\gamma\}$.

We say that $\tau$ is *asymptotic in the future direction (resp. past direction)* to $\gamma$ if $d^{SM}(\dot\tau(t), T^{\pm}\gamma)$ converges to 0 for $t \to +\infty$ (resp. for $t \to -\infty$).

We then obtain the asymptotic version of Question 1.1:

**Question 1.10.** *Let $(M, g)$ be a closed connected Riemannian manifold. What is the minimal number $\mu = \mu(M, g)$ of minimal geodesics $\gamma_1, \ldots, \gamma_\mu$ on $M$ such that every other minimal geodesic is asymptotic in every direction to one of the $\gamma_i$ ? We set $\mu(M, g) := \infty$ if such a finite set of minimal geodesics does not exist.*

The proof of Bangert's lower estimate for the minimal number of minimal geodesics [6, Theorem 4.4] immediately gives the stronger statement

$$\mu(M, g) \geq \dim_{\mathbb{R}} H_1(M;\mathbb{R}).$$

For the Hedlund examples, the equality $\mu(M,g) = \dim_{\mathbb{R}} H_1(M;\mathbb{R})$ is known to be attained in the following cases:

- for $M = T^3$ [19, Section 9],
- for $M = T^n$ with $n \geq 3$ [6, Section 5],
- for compact quotients of Heisenberg groups [1].

We conjecture that on every manifold $M$ with virtually nilpotent fundamental group a Hedlund type construction yields a Riemannian metric $g$ with $\mu(M,g) = \dim_{\mathbb{R}} H_1(M;\mathbb{R})$; see [2] for further discussion. The construction of such a metric on $T^3$ also seems to be the intended goal of the previously mentioned exercise by Burago, Burago, and Ivanov [12, Exercise 8.5.16].

Summarizing we see that for fundamental groups "close" to abelian groups, the Hedlund examples show that Bangert's bounds are optimal in the asymptotic sense. In contrast, one can show that for non-elementary Gromov-hyperbolic fundamental groups we always have $\mu(M,g) = \infty$. For example, on closed manifolds $(M,g)$ with nonpositive sectional curvature, all geodesics (defined on $\mathbb{R}$) are minimal, and for different and non-antipodal unit vectors $v$, $w$ with the same base point, the geodesics $t \mapsto \exp(tv)$ and $t \mapsto \exp(tw)$ are not asymptotic to each other.

**Organisation of this paper.** Section 2 contains preliminaries on minimal geodesics, geometric Hurewicz maps, and the Jacobi map, and defines $\mathbb{R}$- and $\mathbb{Z}$-homologically minimal geodesics.

We recall basics on the stable norm in Section 3. In Section 4, we show that the stable norm and the minimal length have finite distance, a result that essentially follows from D. Burago's work [11], but that does not seem to be worked out in the literature.

Homological asymptotes are discussed in Section 5. The set of terminal/initial asymptotes will allow us to distinguish different types of minimal geodesics, e. g., homologically homoclinic and homologically heteroclinic ones.

Section 6 contains the refined constructions of minimal geodesics. The counting of vertices/edges in centrally symmetric polytopes is given in Section 7. Section 8 contains the proofs of Theorem 1.3, Theorem 1.6, Theorem 1.7, and Theorem 1.8. A variety of examples in dimension 2 and constructions of Hedlund type metrics are collected in Section 9. We also include in this section a comparison to work by Bolotin and Rabinowitz.

For the sake of completeness, we recall basics on minimal geodesics in Appendix A and we spell out the relation between the stable norm and the mass norm on $H_1(\,\bullet\,;\mathbb{R})$ in Appendix B. In Appendix C we summarize additional well-known facts about minimal geodesics on surfaces; we do not intend to include this in the published version.

Readers interested only in the basic Theorem 1.3 need not read the full article: After familiarizing themselves with the basic concepts and results in Sections 2 and 3, they should read Section 4, Subsections 5.1, 5.2 and 5.3, Section 6 until Corollary 6.7, Proposition 7.1 and finally the proof of Theorem 1.3 in Subsection 8.1.

**Acknowledgements.** Bernd Ammann thanks Victor Bangert for bringing his attention to this problem and for many interesting discussions.

## 2. Minimal geodesics, curves, and homology

Before we recall the notion of minimal geodesics, the construction of geometric Hurewicz maps on curves, and the Jacobi map, we will fix some conventions.

2.1. **Conventions.** We use $\mathbb{N} := \{1, 2, \dots\}$ and $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$. All manifolds in this article are assumed to be non-empty.

Geodesics are always assumed to be parametrized by arclength. We use the convention that *a curve in a (smooth) manifold* $M$ is a smooth map $\gamma : I \to M$ from an interval $I$ (usually of positive length) to $M$. An exeception from this notion is the notion of *piecewise smooth curves*, which are defined as continuous maps $\gamma : I \to M$ for which finitely many real numbers $a_1 < a_2 < \ldots < a_k$ with $k \in \mathbb{N}_0$ exist such that we have the following, using $a_0 \coloneqq -\infty$ and $a_{k+1} \coloneqq +\infty$:

$$\text{for all } i \in \{0, 1, \dots, k\}, \text{ the curve } \gamma_{I \cap [a_i, a_{i+1}]} \text{ is smooth.}$$

In particular, every curve is a piecewise smooth curve, but not vice versa. We use the abbreviation "$C^\infty_{\mathrm{pw}}$-curve" for piecewise smooth curves. If – in exceptional cases – curves may have lower regularity (e. g., just continuous), this will be mentioned explicitly. A curve $\gamma$ is *regular* if $\dot\gamma(t) \neq 0$ for all $t$ in the domain.

A $C^\infty_{\mathrm{pw}}$*-loop* in $M$ is a $C^\infty_{\mathrm{pw}}$-path $\gamma : [a, b] \to M$ with $\gamma(a) = \gamma(b)$. A *loop* is defined as a smooth $C^\infty_{\mathrm{pw}}$-loop. Note that for a loop we do not require $\dot\gamma(a) = \dot\gamma(b)$. A *geodesic loop* is a loop that is a geodesic. A *closed curve* is a smooth loop $\gamma\colon [a, b] \to M$ that extends to a (smooth) periodic curve $\gamma\colon \mathbb{R} \to M$ of period length $b - a$. We will often view them as smooth maps $S^1 \to M$. A closed curve is called *simple* if $\gamma\big|_{[a,b]}$ is injective. If a closed curve $\gamma\colon [a, b] \to M$ is a geodesic, it is called a *closed geodesic*, and this is equivalent to saying, that $\gamma$ is a geodesic loop with $\dot\gamma(a) = \dot\gamma(b)$. A *broken geodesic* is defined as a $C^\infty_{\mathrm{pw}}$-curve whose smooth pieces are geodesics.

All differential forms are assumed to be smooth. In a normed space the "unit ball" always denotes the *closed* ball of radius 1 centered in 0.

2.2. **Based and free homotopies of loops.** Let $X$ be a path-connected topological space and let $x_0 \in X$. As a set, $\pi_1(X, x_0)$ is the set $[(S^1, 1), (X, x_0)]_*$ of (continuous) loops in $X$ based at $x_0$ modulo pointed (continuous) homotopies. We will denote the based homotopy class of a based loop $\gamma$ by $[\gamma]_*$.

For geometric considerations it is often better to work with *free homotopies* of loops instead. Let $[S^1, X]$ be the set of (continuous) loops in $X$ modulo (continuous) homotopies; in this case, we do not require constraints on the basepoint – neither for the loops, nor for the homotopies. The free homotopy class of a loop $\gamma$ is denoted by $[\gamma]$. Forgetting the basepoint defines a well-defined map $\mathrm{forget}_{x_0} : \pi_1(X, x_0) \to [S^1, X]$, given by $[\gamma]_* \mapsto [\gamma]$. As $X$ is path-connected, the map $\mathrm{forget}_{x_0}$ is surjective.

Conjugacy classes of $\pi_1(X, x_0)$ will sometimes be called $\pi_1$*-conjugacy classes*. The $\mathrm{forget}_{x_0}$-preimages of singletons in $[S^1, X]$ are precisely the $\pi_1$-conjugacy classes in $\pi_1(X, x_0)$. If $\gamma$ is a $C^0$-loop, then $(\mathrm{forget}_{x_0})^{-1}(\{[\gamma]\})$ is the $\pi_1$*-conjugacy class represented by* $\gamma$.

If $X$ is a connected smooth manifold, then we may replace continuous curves and homotopies by curves and homotopies of regularity $C^k$, $C^\infty$ or $C^\infty_{\mathrm{pw}}$ without changing the results.

2.3. **Minimal geodesics.** Minimal geodesics are Riemannian geodesic lines that lift to "metric" geodesic lines on the Riemannian universal covering:

**Definition 2.1** (minimizing/minimal geodesic)**.** Let $M$ be a closed connected Riemannian manifold and let $\widetilde{M}$ be its Riemannian universal covering.

- A (Riemannian) geodesic $\gamma\colon I \to \widetilde{M}$ on $\widetilde{M}$, defined on a closed interval $I \subset \mathbb{R}$ and parametrized by arclength, is *minimizing* if it is metrically isometric, i.e., if
$$\forall_{s,t\in I} \quad d\bigl(\gamma(t), \gamma(s)\bigr) = |t - s|$$
holds, where $d\colon \widetilde{M} \times \widetilde{M} \to \mathbb{R}_{\geq 0}$ denotes the metric induced by the Riemannian metric on $\widetilde{M}$.
- A geodesic $\gamma\colon I \to M$ on $M$ is *minimal* if $I = \mathbb{R}$ and one (whence every) lift of $\gamma$ to $\widetilde{M}$ is minimizing.
- Geodesics on $M$ are *geometrically equivalent* if there exists a reparametrization that transforms one into the other. Otherwise they are called *geometrically distinct*.

**Example 2.2.** Let $M := S^1 \times N$, where $(N, h)$ is a simply connected closed Riemannian manifold. We view $S^1$ as the unit circle in $\mathbb{C}$, parametrized by $t \mapsto \exp(it)$. We choose a smooth function $f\colon N \to [1, \infty)$ that attains its minimal value 1 only in single point $p \in N$. We define the Riemannian metric $g = f^2 dt^2 + h$ on $M$, i.e., the warped product metric with warping function $f$. One easily sees that $t \mapsto \bigl(\exp(it), p\bigr)$ is a minimal geodesic in $(M, g)$, and that this is up to geometric equivalence the only one. By construction, $\dim_{\mathbb{R}} H_1(M; \mathbb{R}) = 1$.

We will construct minimal geodesics with the following lemma:

**Lemma 2.3** (limiting geodesics)**.** *Let $M$ be a closed connected Riemannian manifold, and let $\widehat{M} \to M$ be a Riemannian covering. Let $(\sigma_i\colon [a_i, b_i] \to \widehat{M})_{i\in\mathbb{N}}$ be a sequence of minimizing geodesics on $\widehat{M}$ with the following properties:*

- *We have $\lim_{i\to\infty} a_i = -\infty$ and $\lim_{i\to\infty} b_i = \infty$.*
- *The sequence $(\dot{\sigma}_i(0))_{i\in\mathbb{N}}$ converges to some $v_\infty \in T\widehat{M}$.*

*Then*
$$\begin{aligned} \sigma_\infty\colon \mathbb{R} &\to \widehat{M} \\ t &\mapsto \exp_{\widehat{M}}(t \cdot v_\infty) \end{aligned}$$
*is a minimizing geodesic and the $(\sigma_i)_{i\in\mathbb{N}}$ converge on each compact interval uniformly (in the $C^\infty$-topology) to $\sigma_\infty$.*

A proof is provided in Appendix A.

2.4. **Curves and homology.** The geometric Hurewicz map is defined via integration. In contrast with the topological Hurewicz map for continuous *loops*, the geometric Hurewicz map depends on how we represent de Rham cohomology classes by forms. However, this dependence will have no relevance for our following arguments. By $H_1(M; \mathbb{Z})_{\mathbb{R}} \subset H_1(M; \mathbb{R})$, we denote the integral part of $H_1(M; \mathbb{R})$, i.e., the image of $H_1(M; \mathbb{Z})$ under the change-of-coefficients map from $\mathbb{Z}$ to $\mathbb{R}$. Then $H_1(M; \mathbb{Z})_{\mathbb{R}}$ is a lattice in $H_1(M; \mathbb{R})$.

**Setup 2.4.** Let $M$ be a closed connected Riemannian manifold and let $b := \dim_{\mathbb{R}} H_1(M; \mathbb{R})$. Let $(\beta_1, \dots, \beta_b)$ be a $\mathbb{Z}$-module basis of $H_1(M; \mathbb{Z})_{\mathbb{R}}$, which is also an $\mathbb{R}$-basis of $H_1(M; \mathbb{R})$. Let $\alpha^1, \dots, \alpha^b$ be closed 1-forms on $M$ such that

$([\alpha^1], \dots, [\alpha^b])$ is the basis of $H^1_{\mathrm{dR}}(M)$ dual to $(\beta_1, \dots, \beta_b)$, viewed as a vector space basis for $H_1(M;\mathbb{R})$.

**Definition 2.5** (geometric Hurewicz map)**.** In the situation of Setup 2.4, for each $C^\infty_{\mathrm{pw}}$-curve $\gamma\colon I \to M$, we let $h(\gamma)$ denote the unique element of $H_1(M;\mathbb{R})$ with

$$\forall_{j\in\{1,\dots,b\}} \quad \int_\gamma \alpha^j = \bigl\langle [\alpha^j], h(\gamma) \bigr\rangle \,.$$

Here, $\langle \,\cdot\,, \cdot\, \rangle$ denotes the integration pairing between $H^1_{\mathrm{dR}}(M)$ and $H_1(M;\mathbb{R})$.

**Remark 2.6** (additivity)**.** In the situation of Setup 2.4, using the additivity of integration along curves, we obtain: If $\gamma$ and $\gamma'$ are $C^\infty_{\mathrm{pw}}$-curves on $M$ such that the endpoint of $\gamma$ is the start point of $\gamma'$, then

$$h(\gamma * \gamma') = h(\gamma) + h(\gamma') \quad \text{and} \quad h(\overline{\gamma}) = -h(\gamma)\,,$$

where "$*$" denotes the concatenation of curves (if well-defined) and "$^-$" denotes the orientation reversal of curves.

**Remark 2.7** (loops)**.** In the situation of Setup 2.4, the integration duality between $H_1(M;\mathbb{R})$ and $H^1_{\mathrm{dR}}(M)$ shows: If $\gamma$ is a $C^\infty_{\mathrm{pw}}$-loop in $M$, then $h(\gamma)$ is the singular homology class represented by the loop $\gamma$; in particular, $h(\gamma)$ then lies in $H_1(M;\mathbb{Z})_{\mathbb{R}}$. Furthermore, we have for every closed 1-form $\eta \in \Omega^1(M)$:

$$\int_\gamma \eta = \bigl\langle [\eta], h(\gamma) \bigr\rangle \,.$$

**Remark 2.8.** Again we assume the situation of Setup 2.4. If $\eta \in \Omega^1(M)$ is a linear combination of $\alpha^1, \dots, \alpha^b$, then we have for every $C^\infty_{\mathrm{pw}}$-curve $\gamma\colon [a,b] \to M$:

$$\int_\gamma \eta = \bigl\langle [\eta], h(\gamma) \bigr\rangle \,.$$

However, in general, this property does *not* hold for arbitrary closed 1-forms $\eta$ when $\gamma$ is not a loop. For example, let $f : M \to \mathbb{R}$ be smooth with $f(\gamma(a)) = 0$ and $f(\gamma(b)) = 1$, and let $\eta \coloneqq df$. Then, $[\eta] = 0$ but

$$\int_\gamma \eta = 1 \neq 0 = \bigl\langle [\eta], h(\gamma) \bigr\rangle \,.$$

**Definition 2.9** (Jacobi map)**.** In the situation of Setup 2.4, in addition, we pick a basepoint $x_0$ in $M$ and a lift $\widetilde{x}_0 \in \widetilde{M}$ of $x_0$. For $j \in \{1, \dots, b\}$, let $\widetilde{\alpha}^j$ denote the pull-back of $\alpha^j$ to $\widetilde{M}$. Then, there is a unique smooth function $J^j\colon \widetilde{M} \to \mathbb{R}$ with

$$\mathrm{d}J^j = \widetilde{\alpha}^j \quad \text{and} \quad J^j(\widetilde{x}_0) = 0\,.$$

The *Jacobi map* is defined as the sum

$$J \coloneqq \sum_{j=1}^b J^j \cdot \beta_j \colon \widetilde{M} \to H_1(M;\mathbb{R})\,.$$

**Remark 2.10.** In the situation of Definition 2.9, the Jacobi map $J\colon \widetilde{M} \to H_1(M;\mathbb{R})$ is smooth. Moreover, the Jacobi map is related to the geometric Hurewicz map as follows: If $\gamma\colon [a,b] \to M$ is a smooth curve and $\widetilde{\gamma}\colon [a,b] \to \widetilde{M}$ is a lift to $\widetilde{M}$, then

$$h(\gamma) = J\bigl(\widetilde{\gamma}(b)\bigr) - J\bigl(\widetilde{\gamma}(a)\bigr)\,.$$

Indeed, for all $j \in \{1, \ldots, b\}$, we have

$$\begin{aligned}\big\langle [\alpha^j], h(\gamma) \big\rangle &= \int_\gamma \alpha^j = \int_{\tilde\gamma} \widetilde{\alpha}^j = \int_{\tilde\gamma} dJ_j = J_j\big(\widetilde{\gamma}(b)\big) - J_j\big(\widetilde{\gamma}(a)\big) \\ &= \Big\langle [\alpha^j], J\big(\widetilde{\gamma}(b)\big) - J\big(\widetilde{\gamma}(a)\big) \Big\rangle .\end{aligned}$$

**Remark 2.11.** It follows from Remark 2.7 and the additivity in Remark 2.6 that the Jacobi map $J\colon \widetilde{M} \to H_1(M;\mathbb{R})$ descends to to a map $\not{J} : M \to H_1(M;\mathbb{R})/H_1(M;\mathbb{Z})_\mathbb{R}$. The map $\not{J}$ is Gromov's "Jacobi mapping" [18, 4.21] and $H_1(M;\mathbb{R})/H_1(M;\mathbb{Z})_\mathbb{R}$ is called the "Jacobi variety"; literally, only the case that $H_1(M;\mathbb{Z})$ is torsionfree is treated, but the extension to the general case is straightforward.

2.5. **$\mathbb{R}$-homologically minimal and $\mathbb{Z}$-homologically minimal geodesics.** We will now introduce two stronger versions of minimality for geodesics.

For this purpose let $\Gamma := \pi_1(M)$ be the fundamental group of a closed connected Riemannian manifold $(M, g)$ and let $[\Gamma, \Gamma]$ be its commutator subgroup. The Hurewicz map factors to an isomorphism $\Gamma/[\Gamma,\Gamma] \to H_1(M;\mathbb{Z})$. Thus $\widehat{M}^\mathbb{Z} := \widetilde{M}/[\Gamma,\Gamma]$ is a normal covering of $M$ with deck transformation group (canonically isomorphic to) $H_1(M;\mathbb{Z})$. Similarly, if $T$ is the torsion subgroup of $H_1(M;\mathbb{Z})$, then the change-of-coefficients map $H_1(M;\mathbb{Z}) \to H_1(M;\mathbb{R})$ induces an isomorphism $H_1(M;\mathbb{Z})/T \to H_1(M;\mathbb{Z})_\mathbb{R}$. The deck transformation group of the normal covering $\widehat{M}^\mathbb{R} := \widehat{M}^\mathbb{Z}/T$ of $M$ is (canonically isomorphic to) $H_1(M;\mathbb{Z})_\mathbb{R}$. The Jacobi map $J\colon \widetilde{M} \to H_1(M;\mathbb{R})$ factors to maps $\widehat{J}^\mathbb{Z}\colon \widehat{M}^\mathbb{Z} \to H_1(M;\mathbb{R})$ and $\widehat{J}^\mathbb{R}\colon \widehat{M}^\mathbb{R} \to H_1(M;\mathbb{R})$. We thus obtain the following sequence of covering maps:

$$\widetilde{M} \to \widehat{M}^\mathbb{Z} \to \widehat{M}^\mathbb{R} \to M .$$

**Definition 2.12** ($\mathbb{R}$-homologically and $\mathbb{Z}$-homologically minimal geodescis)**.** Let $M$ be a closed connected Riemannian manifold.

- A geodesic $\gamma\colon \mathbb{R} \to M$ on $M$ is $\mathbb{Z}$-*homologically minimal* if one (whence every) lift of $\gamma$ to $\widehat{M}^\mathbb{Z}$ is minimizing.
- A geodesic $\gamma\colon \mathbb{R} \to M$ on $M$ is $\mathbb{R}$-*homologically minimal* if one (whence every) lift of $\gamma$ to $\widehat{M}^\mathbb{R}$ is minimizing.

For a geodesic $\mathbb{R} \to M$, "$\mathbb{R}$-homologically minimal" implies "$\mathbb{Z}$-homologically minimal", and the latter one implies "minimal". The minimal geodesics constructed in Theorem 1.6 (and thus also the ones in Theorem 1.3) are in fact "$\mathbb{R}$-homologically minimal". Also, the slightly larger class of $\mathbb{Z}$-homologically minimal geodesics has many properties of minimal geodesics, especially the ones related to the stable norm, hold in this larger class (see e. g., Proposition 5.5).

**Comparison to the literature.** Gromov calls $\widehat{M}^\mathbb{R}$ the *abelian covering* and denotes it by $\widetilde{M}_{\mathrm{Ab}}$ [18, $4.22_+$-$4.24_+$]. Bangert considers $\widehat{M}^\mathbb{Z}$ and denotes it by $\bar{M}$ [6, Sec. 3].

## 3. The stable norm

The stable norm on first homology is the homogenisation of the quasi-norm given by the minimal length of representing curves. We recall this construction and basic estimates for the stable norm on the first real homology group $H_1(M;\mathbb{R})$. Our approach is similar to the one of Bangert [6], which again relies on similar notions and statements by Gromov [18, Chap. 4.C.].

In the next section (Theorem 4.1), we will show that the stable norm $\|\boldsymbol{\cdot}\|_{\mathrm{st}}$, restricted to the integral lattice $H_1(M;\mathbb{Z})_\mathbb{R}$, coincides up to a bounded error with the minimal length quasi-norm $N : H_1(M;\mathbb{Z})_\mathbb{R} \to \mathbb{R}_{\geq 0}$. This approximation result is much stronger than the ones proven by Bangert and Gromov.

3.1. **Quasi-norms and their homogenisation.** We recall basic notions and properties for general quasi-norms on abelian groups.

**Definition 3.1** (quasi-norm)**.** Let $V$ be an abelian group. A *quasi-norm* on $V$ is a map $N\colon V \to \mathbb{R}_{\geq 0}$ with the following properties:

- *Positive definiteness.* We have for all $x \in V$ that $N(x) = 0$ if and only if $x = 0$.
- *Symmetry.* For all $x \in V$, we have $N(-x) = N(x)$.
- *Quasi-triangle inequality.* There exists a $\Delta \in \mathbb{R}_{\geq 0}$ such that for all $x, y \in V$ we have
$$N(x+y) \leq N(x) + N(y) + \Delta\,.$$

There is a least constant $\Delta_{\min}$, satisfying the inequality above for all $x, y \in V$, and this constant will be called the $\triangle$-discrepancy of $N$. Note that $N(0) = 0$ implies $\Delta_{\min} \geq 0$.

We will occasionally need further properties of quasi-norms, and more generally of maps $N\colon V \to \mathbb{R}_{\geq 0}$, where $V$ is a $K$-module with $K \in \{\mathbb{N}_0, \mathbb{Z}, \mathbb{R}\}$:

- A map $N\colon V \to \mathbb{R}_{\geq 0}$ is *homogeneous over* $K$ if we have for all $x \in V$ and all $k \in K$ that
$$N(k \cdot x) = |k| \cdot N(x)\,.$$
- The map $N\colon V \to \mathbb{R}_{\geq 0}$ is called *subhomogeneous (over* $\mathbb{N}$*)*, if we have for all $x \in V$ and all $n \in \mathbb{N}$ that
$$N(k \cdot x) \leq |k| \cdot N(x)\,.$$
- The map $N\colon V \to \mathbb{R}_{\geq 0}$ satisfies the *doubling property*, if there exists a $D \in \mathbb{R}_{\geq 0}$ such that for all $x \in V$ we have
$$2 \cdot N(x) \leq N(2 \cdot x) + D\,.$$
Again, there is a least constant $D_{\min}$ for which this inequality holds for all $x \in V$. This least constant will be called the *doubling constant of* $N$.

**Proposition 3.2** (homogenisation of quasi-norms)**.** *Let $N\colon V \to \mathbb{R}_{\geq 0}$ be a quasi-norm with $\triangle$-discrepancy $\Delta$ on the abelian group $V$. For each $x \in V$, the limit*
$$\|x\| := \lim_{k\to\infty} \frac{N(k \cdot x)}{k}$$
*exists and we have*
$$\|x\| = \inf_{k\in\mathbb{N}} \frac{N(k \cdot x) + \Delta}{k}\,.$$
*The map $\|\boldsymbol{\cdot}\| : V \to \mathbb{R}_{\geq 0}$ is homogeneous over $\mathbb{Z}$ and satisfies the triangle inequality; we call $\|\boldsymbol{\cdot}\|$ the* homogenisation of $N$*. If, additionally, $N$ is subhomogeneous, then*
$$\forall_{x\in V} \quad \|x\| \leq N(x)\,.$$

One may easily construct examples for which $\|\boldsymbol{\cdot}\|$ is no longer positive definite. Positive definiteness is lost, e.g., on all non-trivial torsion elements of $V$.

*Proof.* Let $x \in V$. The sequence $(s_k)_{k\in\mathbb{N}}$ with $s_k := N(k\cdot x)+\Delta$ is subadditive, i. e., $s_{k+\ell} \leq s_k + s_\ell$ for all $k,\ell \in \mathbb{N}$. By the Fekete lemma [31, § I.3.1], the limit $\lim_{k\to\infty}(N(k\cdot x)+\Delta)/k$ exists and

$$\lim_{k\to\infty} \frac{N(k\cdot x)+\Delta}{k} = \inf_{k\in\mathbb{N}} \frac{N(k\cdot x)+\Delta}{k}\,.$$

In particular, also $\lim_{k\to\infty} N(k\cdot x)/k$ exists and

$$\lim_{k\to\infty} \frac{N(k\cdot x)}{k} = \inf_{k\in\mathbb{N}} \frac{N(k\cdot x)+\Delta}{k}\,.$$

By construction, $\|\bullet\|$ is homogeneous over $\mathbb{N}$ and $\|\bullet\|$ inherits symmetry from $N$. Therefore, $\|\bullet\|$ is homogeneous over $\mathbb{Z}$.

For the triangle inequality, let $x,y \in V$. Then, the quasi-triangle inequality for $N$ shows that

$$\begin{aligned}
\|x+y\| &= \lim_{k\to\infty} \frac{N(k\cdot(x+y))}{k} \\
&\leq \lim_{k\to\infty} \frac{N(k\cdot x)+N(k\cdot y)+\Delta}{k} = \lim_{k\to\infty}\frac{N(k\cdot x)}{k} + \lim_{n\to\infty}\frac{N(k\cdot y)}{k} + 0 \\
&= \|x\| + \|y\|\,.
\end{aligned}$$

Finally, let $N$ be subhomogeneous. Then,

$$\|x\| = \lim_{k\to\infty}\frac{N(k\cdot x)}{k} \leq \lim_{k\to\infty}\frac{|k|\cdot N(x)}{k} = N(x)\,,$$

as claimed. □

**Proposition 3.3.** *Let $N\colon V\to\mathbb{R}_{\geq 0}$ be a quasi-norm on the abelian group $V$ satisfying the doubling property with doubling constant $D$. Then the homogenisation $\|\bullet\|$ of $N$ satisfies*

$$\forall_{x\in V}\quad N(x) \leq \|x\| + D\,.$$

*Proof.* Let $x \in V$. From the doubling property, we inductively obtain

$$\forall_{k\in\mathbb{N}}\quad 2^k\cdot N(x) \leq N(2^k\cdot x) + \sum_{j=0}^{k-1} 2^j\cdot D \leq N(2^k\cdot x) + 2^k\cdot D\,.$$

Taking the limit over the subsequence of binary powers shows therefore that

$$\begin{aligned}
N(x) &= \lim_{k\to\infty}\frac{2^k\cdot N(x)}{2^k} \leq \lim_{k\to\infty}\frac{N(2^k\cdot x)+2^k\cdot D}{2^k} \\
&= \|x\| + D\,.
\end{aligned}$$

□

For the sake of completeness, we record a discrete example:

**Remark 3.4** (word lengths on finitely generated abelian groups [12, Exc. 8.5.8])**.** Let $V$ be a finitely generated free abelian group, let $S\subset V$ be a finite generating set of $V$, and let $N$ be the word norm on $V$ associated with $S$. Then $N$ is subhomogeneous and $N$ has the doubling property: Indeed, let $k := \#S$ and $S = \{s_1,\dots,s_k\}$. For $x\in V$ we choose an $N$-minimal $S$-representation $2\cdot x = \sum_{j=1}^k a_j\cdot s_j$ of $2\cdot x$ with $a_1,\dots,a_k\in\mathbb{Z}$. We write $a_j = 2\cdot b_j + r_j$ with $b_j\in\mathbb{Z}$ and $r_j\in\{0,1\}$ and consider the quasi-half

$$x' := \sum_{j=1}^k b_j\cdot s_j$$

of $2 \cdot x$. We obtain

$$\begin{aligned} 2 \cdot N(x) &\leq 2 \cdot N(x') + 2 \cdot N(x - x') \\ &\leq 2 \cdot \sum_{j=1}^{k} |b_j| + 2 \cdot k = \sum_{j=1}^{k} |a_j - r_j| + 2 \cdot k \leq \sum_{j=1}^{k} |a_j| + k + 2 \cdot k \\ &= N(2 \cdot x) + 3 \cdot \#S\,, \end{aligned}$$

which proves that $N$ has the doubling property. Proposition 3.3 and Proposition 3.2 show that $N$ and its homogenisation have uniformly finite distance on $V$.

This argument can be extended to word norms on all finitely generated abelian groups. However, in contrast with the homogenisation process for quasi-morphisms, word norms on general amenable groups need not have finite distance to their homogenisation: For example, the centre of the integral 3-dimensional Heisenberg group is distorted with respect to word norms.

3.2. **The minimal length quasi-norm.** We first fix notation for the minimal length of representing curves, which defines a quasi-norm, and then introduce the stable norm via homogenisation and extension.

**Definition 3.5.** Let $M$ be a closed connected Riemannian manifold. We define the *minimal length map* as the function

$$\begin{aligned} N \colon H_1(M;\mathbb{Z})_{\mathbb{R}} &\to \mathbb{R}_{\geq 0} \\ x &\mapsto \inf\bigl\{\mathcal{L}(\gamma) \bigm| \gamma \text{ is a loop on } M \text{ with } h(\gamma) = x\bigr\}\,, \end{aligned}$$

where $\mathcal{L}(\gamma)$ denotes the length of the curve $\gamma$.

**Remark 3.6.** As this definition is based on loops, only the canonical part of the Hurewicz map is used. In the definition, we used our convention that loops are supposed to be smooth. However, let us explain why the value $N(x)$ remains unchanged if we weaken this condition to $C^\infty_{\mathrm{pw}}$-curves, to (piecewise) $C^1$-curves, or even to $C^0$-curves. It follows from standard smoothing techniques for loops e. g., [26, §16 and 17] that the inclusion of the space of smooth loops into the space curves of regularity $C^\infty_{\mathrm{pw}}$, (piecewise) $C^1$ or $C^0$ is a homotopy equivalence if all these spaces are equipped with their natural topologies. The homotopy inverses can be constructed by approximations by broken geodesics and rounding off the corners. These constructions can be performed in such a way that they do not increase lengths; therefore, the definition of $N$ does not depend on the choice of curve regularity.

Furthermore, the infimum in the definition of $N$ is attained (see Lemma A.3). If the infimum is attained in some curve $\gamma$ and if the class $[\gamma] \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ is non-zero, then $\gamma$ can be reparametrized (preserving the orientation) to a closed geodesic. Such a closed geodesic may be parametrized by arclength as $\gamma : [a, a + N(x)] \to M$ for every $a \in \mathbb{R}$. In the case $[\gamma] = 0$, the curve $\gamma$ is then constant. This implies that *groupnorm* is positive definite.

In particular: If $\gamma\colon I \to M$ is a $C^1$-loop with $\mathcal{L}(\gamma) = N\bigl(h(\gamma)\bigr)$, then every lift $I \to \widehat{M}^{\mathbb{R}}$ of $\gamma$ to $\widehat{M}^{\mathbb{R}}$ is minimizing. Hence, sequences of such geodesics can be used to construct minimal geodesics on $M$ via Lemma 2.3. This will be the starting point for the constructions in the proof of Theorem 1.3.

**Comparison to the literature.** The map $N$ above is denoted as $f$ by Bangert [6].

Gromov does not treat the norm $N$, but a variant called "length" defined on $H_1(M;\mathbb{Z})$ instead of $H_1(M;\mathbb{Z})_\mathbb{R}$, using loops in a fixed basepoint $v_0 \in M$ [18, Chap. 4.C.]. More precisely, for every class in $\pi_1(M, v_0)$, Gromov considers the infimum (= minimum) of lengths of loops based at $v_0$ representing this given class, leading to a function $\pi_1(M, v_0) \to \mathbb{R}$. By minimizing this function over preimages of the Hurewicz map $\pi_1(M, v_0) \to H_1(M;\mathbb{Z})$, Gromov obtains a map length: $H_1(M;\mathbb{Z}) \to \mathbb{R}_{\geq 0}$ that is positive definite, symmetric, satisfies the triangle inequality (i. e., $\Delta_{\min} = 0$ in the quasi-triangle inequality), and is subhomogeneous. There is a constant $C$ with the property: if $\bar{\alpha}$ denotes the image of $\alpha \in H_1(M;\mathbb{Z})$ in $H_1(M;\mathbb{Z})_\mathbb{R}$, then

$$N(\bar{\alpha}) \leq \operatorname{length}(\alpha) \leq N(\bar{\alpha}) + C\,.$$

Then $\|\bar{\alpha}\|_{\mathrm{st}}$ (in our notation) coincides with $\|\alpha_\mathbb{R}\|$ in Gromov's notation; however, these two norms are slightly differently introduced.

3.3. **The stable norm on first homology.** We introduce the stable norm on $H_1(M;\mathbb{R})$ as the homogenization of the minimal length map $N : H_1(M;\mathbb{Z})_\mathbb{R} \to \mathbb{R}_{\geq 0}$. Our approach to the stable norm is close to Bangert's [6, Sec. 2]. It can be shown that the norm defined this way is equivalent to the classical stable norm by Federer [14, §3], also called the *mass* of a real homology class [18, 4.15–4.17]; see Definition B.1 for $K = \mathbb{R}$ for a definition in easier notation. The equivalence follows from Gromov's work [18, 4.18 and $4.20\frac{1}{2}\mathrm{bis}_+$] by using results from Federer [14, §5]. For the sake of being self-contained we present the proof of this equivalence in Appendix B.

**Proposition 3.7** (the stable norm)**.** *Let $M$ be a closed connected Riemannian manifold. Then:*

(1) *The minimal length map $N\colon H_1(M;\mathbb{Z})_\mathbb{R} \to \mathbb{R}_{\geq 0}$ from Definition 3.5 is a subhomogeneous quasi-norm on $H_1(M;\mathbb{Z})_\mathbb{R}$ in the sense of Definition 3.1 (with $\triangle$-discrepancy at most $2 \cdot \operatorname{diam} M$). We refer to $N$ as the* minimal length quasi-norm*.*
(2) *The homogenisation (in the sense of Proposition 3.2) of $N$ on $H_1(M;\mathbb{Z})_\mathbb{R}$ extends uniquely to a norm $\|\bullet\|_{\mathrm{st}}$ on $H_1(M;\mathbb{R})$, the* stable norm*.*

The minimal length quasi-norm also satisfies a doubling property, as we will see in Proposition 4.5.

*Proof.* Ad (1): This can be checked on representatives, using the flexibility in regularity provided by Remark 3.6; for the quasi-triangle inequality, we connect short loops by paths of length at most $\operatorname{diam} M$.

Ad (2): We define $\|\bullet\|_{\mathrm{st}} : H_1(M;\mathbb{R}) \to \mathbb{R}_{\geq 0}$ as the homogenisation of $N$, as described in Proposition 3.2. In particular, $\|\bullet\|_{\mathrm{st}}$ is homogeneous and satisfies the triangle inequality. Symmetry follows from the symmetry of $N$. It remains to prove $\|x\|_{\mathrm{st}} > 0$ for all $x \in H_1(M;\mathbb{R}) \setminus \{0\}$.

We fix an inner product $\langle \bullet, \bullet \rangle_{H_1}$ on $H_1(M;\mathbb{R})$, which induces a norm $\|\bullet\|_{H_1}$, and thus a length functional $\mathcal{L}^{H_1}$ for curves. Let $J\colon \widetilde{M} \to H_1(M;\mathbb{R})$ be the Jacobi map (Definition 2.9). This map is differentiable with $\mathrm{d}J = \sum_{i=1}^{b} \alpha^i \beta_i \in \Omega^1(M) \otimes H_1(M;\mathbb{R})$. Let $\left\|\mathrm{d}J|_p\right\|_{\mathrm{Op}}$ denote the operator norm of

$$\mathrm{d}J|_p \colon (T_pM, g_p) \to (H_1(M;\mathbb{R}), \langle \bullet, \bullet \rangle_{H_1})\,.$$

As $\mathrm{d}J|_p$ depends continuously on $p$ and as it is invariant under the deck transformation group, which acts cocompactly, we see that the supremum in

$$\mathcal{K} \coloneqq \sup_{p \in \widetilde{M}} \left\| \mathrm{d}J|_p \right\|_{\mathrm{Op}}$$

is attained and thus finite. For every finite $C^\infty_{\mathrm{pw}}$-curve $\gamma$, we get

$$\mathcal{L}^{H_1}(J \circ \gamma) \leq \mathcal{K} \cdot \mathcal{L}^g(\gamma)\,.$$

This immediately implies that for every $x \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ we obtain

$$\|x\|_{H_1} \leq \mathcal{K} \cdot N(x)\,.$$

We apply homogenisation, which does not affect $\|\bullet\|_{H_1}$ as it is already homogeneous, and get

$$\frac{1}{\mathcal{K}} \cdot \|x\|_{H_1} \leq \|x\|_{\mathrm{st}}\ .$$

Positive definiteness on the left-hand side implies positive definiteness on the right-hand side. ☐

**Example 3.8.** Let $M = T^n$ be an $n$-dimensional torus, which we will equip with two kinds of metrics $g$. We identify $T^n$ with $\mathbb{R}^n/\Gamma$ where $\Gamma$ is a lattice, i. e., a discrete cocompact subgroup of $\mathbb{R}^n$. In both cases we use the notation of Setup 2.4 and the construction of the Jacobi map in Definition 2.9. It is convenient to choose the basepoint $x_0 = 0$ and the forms $\alpha^i = dx^i$, where $x = (x^1, \dots, x^n)$ are the standard coordinates of $\widetilde{M} = \mathbb{R}^n$. Then we get $J^i(x) = x^i$, i. e., $J$ corresponds to the identity map under the canonical isomorphism $H_1(T^n;\mathbb{R}) \cong \mathbb{R}^n$. Moreover, we have a compatible canonical identifications $H_1(T^n;\mathbb{Z}) \cong H_1(M;\mathbb{Z})_{\mathbb{R}} \cong \Gamma$.

(1) If $g$ is flat, then we can identify $(T^n, g)$ with $\mathbb{R}^n/\Gamma$, where $\mathbb{R}^n$ carries the Euclidean metric. Then $N$ and $\|\bullet\|_{\mathrm{st}}$ coincide with the Euclidean norm. The stable norm unit ball is thus the standard ball.

    Every geodesic on $(T^n, g)$ can be extended to $\mathbb{R}$ and this extension lifts to a straight affine line in $\mathbb{R}^n$. Parametrized by arclength, these geodesics are minimal.

(2) Let $n \geq 3$. For the Hedlund metrics $g$ on $T^n$, the stable norm is the $\ell^1$-norm with respect to the chosen basis of $H_1(T^n;\mathbb{R})$ [6, Prop. 5.8] (the argument given by Bangert for $n = 3$ can easily be extended to $n \geq 3$). In particular, the stable norm unit ball has the combinatorial type of the $n$-dimensional cross-polytope.

**Remark 3.9.** If $M$ is a closed connected Riemannian manifold, then the stable norm unit ball $B \subset H_1(M;\mathbb{R})$ is convex, compact, and centrally symmetric. Several types of convex bodies are possible, e. g., polytopes, smooth domains, combinations of flat and smooth boundaries and much more. But not every convex, compact, and centrally symmetric body will arise as the stable norm unit ball for a Riemannian metric, see Remark 3.10.

However, if $\dim M \geq 3$ and $P$ is a convex compact centrally symmetric polyhedron in $H_1(M;\mathbb{R})$ whose vertices have a rational direction (with respect to $H_1(M;\mathbb{Z})_{\mathbb{R}}$), then there exists a Riemannian metric on $M$ whose stable norm unit ball coincides with $P$. Suitable metrics are provided by the Hedlund metrics, introduced on arbitrary closed connected manifolds of dimension at least 3 in the first author's diploma thesis [1]. The fact that every stable norm unit ball as above may be achieved by such Hedlund metrics follows immediately from the necessary conditions on $\mathbb{R}$-homologically minimal geodesics in this diploma thesis [1, IV.1.e Homolog minimale Geodätische]; in fact, Bangert's minimal geodesics are $\mathbb{R}$-homologically minimal geodesics and thus these methods determine the norm ball of the Hedlund examples [1, Korollar IV.1.16, Bemerkung IV.1.18].

Hedlund metrics on closed connected manifolds were re-introduced later [4, 20] and the statements of this remark have been worked out explicitly in these articles. We will give more details on Hedlund examples in Subsection 9.2.

**Remark 3.10.** For a given closed manifold $M$, one can ask whether there is a Riemannian metric on $M$ such that the given norm is its stable norm. This question is wide open, even for tori. However, some obstruction are known. Points in $H_1(M;\mathbb{R}) \cong H_1(M;\mathbb{Z})_\mathbb{R} \otimes_\mathbb{Z} \mathbb{R} \cong (H_1(M;\mathbb{Z})_\mathbb{R} \otimes_\mathbb{Z}) \otimes_\mathbb{Z} \mathbb{R}$ are called irrational if their coefficients are linearly independent over $\mathbb{Q}$.

On the 2-torus, Bangert [7] has proved that the stable norm is differentiable at every irrational point. If the stable norm is differentiabile in a rational point, then the torus is foliated by minimal geodesics in this direction.

In dimension larger than 2, there is more flexibility to obtain a given norm, but some obstructions remain. On $n$-tori, $n \geq 3$, Burago, Ivanov, and Kleiner [13] have proved that in irrational points, the stable norm has a directional derivative in at least one non-radial direction. This implies that vertices of the stable norm ball have rational directions. In particular if $B$ is a polytope, then all of its vertices lie in rational directions.

3.4. **Some definitions for convex bodies.** Let us recall some standard terminology for convex bodies.

Let us assume that $B$ is a closed convex subset of a finite-dimensional $\mathbb{R}$-vector space $V$ with $0 \in \mathring{B}$. An *exposed face* of $B$ is a non-empty subset $F \subset B$ for which there exists an affine hyperplane $H \subset V$ with

$$H \cap \partial B = F \quad \text{and} \quad H \cap \mathring{B} = \emptyset\,;$$

in this case, $H$ is called a *supporting hyperplane* for the exposed face $F$ of $B$. An *exposed vertex* of $B$ is a point $x \in \partial B$ such that $\{x\}$ is an exposed face of $B$. An *exposed edge* of $B$ is an exposed face of dimension 1.

If $H$ is a supporting hyperplane for $F$, then $0 \notin H$, and thus there is a unique $\omega_H$ in the dual space $V^*$ with $H = \{x \in V \mid \omega_H(x) = 1\}$. Then $F = \{x \in B \mid \omega_H(x) = 1\}$.

It is easy to show that every closed convex subset of $V$ is the intersection of the closed halfspaces containing the given convex subset. As a consequence $\partial B$ is the union of all exposed faces of $B$.

The definition of an exposed point is tightly related to the definition of an *extreme point*. A point $v$ in a convex body $B$ is called *extreme* if it does not lie in any open line segment between two different points of $B$. All exposed points are extreme, but not vice versa.

The usage of the words "extreme" and "exposed" may be considered as confusing: one easily sees that a point $x \in \partial B$ is exposed with supporting hyperplane $H$ if, and only if, it is the unique maximum of the linear function $\omega_H \colon V \to \mathbb{R}$ restricted to $B$. Not all extreme points satisfy this property. In this article, exposed points will play an important role, but the definition of extreme points will only be of marginal importance.

Often we will also abbreviate the term “an exposed edge of $B$” by simply saying “an edge of $B$”, and similarly we will abbreviate “an exposed face” by “a face”. However, one should keep in mind that usually in the literature more general notions of faces and edges are used.

For compact convex polytopes in finite-dimensional $\mathbb{R}$-vector spaces, all faces (and vertices) are exposed.

The main focus of the article lies on closed Riemannian manifolds with a finite number of geometrically distinct minimal geodesics. Using a result by Bangert [6, Theorem 4.4], see also Subsection 5.4, this assumption implies that the stable unit ball in $H^1_{\mathrm{dR}}(M)$ is a polytope. Thus all faces are exposed, and extremality coincides with exposedness in this case.

3.5. **The dual norm.** Finally, we collect elementary properties of the dual norm of the stable norm. In fact, the following considerations apply to all norms on finite-dimensional $\mathbb{R}$-vector spaces. For convenience, we formulate most of the statements directly for the stable norm.

**Definition 3.11** (dual norm)**.** Let $M$ be a closed connected Riemannian manifold. The dual norm of $\|\bullet\|_{\mathrm{st}}$ on $H^1_{\mathrm{dR}}(M)$ is denoted by $\|\bullet\|^*_{\mathrm{st}}$. More explicitly: For all $\omega \in H^1_{\mathrm{dR}}(M)$, we have

$$\|\omega\|^*_{\mathrm{st}} = \sup_{x\in H_1(M;\mathbb{R})\setminus\{0\}} \frac{\big|\langle\omega, x\rangle\big|}{\|x\|_{\mathrm{st}}}\,. \tag{3.1}$$

In view of compactness of $\{x \in H_1(M;\mathbb{R}) \mid \|x\|_{\mathrm{st}} = 1\}$, the supremum is attained (if $H_1(M;\mathbb{R}) \not\cong 0$).

For $\omega \in H^1_{\mathrm{dR}}(M) \setminus \{0\}$ we define the hyperplane

$$H_\omega \coloneqq \big\{x \in H_1(M;\mathbb{R}) \;\big|\; \langle\omega, x\rangle = 1\big\}\,. \tag{3.2}$$

Every hyperplane $H$ with $0 \notin H$ is obtained this way.

Let us assume in this entire subsection that $M$ is a closed connected Riemannian manifold, and that $B \subset H_1(M;\mathbb{R})$ is the stable norm unit ball.

**Lemma 3.12.** *With the definitions above $H_\omega$ is a supporting hyperplane for a face of $B$ if, and only if, $\|\omega\|^*_{\mathrm{st}} = 1$.*

*Proof.* The proof is straigthforward, using that the supremum in (3.1) is attained. □

**Definition 3.13.** Let $M$ be a closed connected Riemannian manifold, let $B \subset H_1(M;\mathbb{R})$ be the stable norm unit ball, and let $\omega \in H^1_{\mathrm{dR}}(M)$ with $\|\omega\|^*_{\mathrm{st}} = 1$. We define the hyperplane $H_\omega$ as in (3.2). Equation 3.1 and $\|\omega\|^*_{\mathrm{st}} = 1$ imply $H_\omega \cap \mathring{B} = \emptyset$. The *exposed face of $B$ defined by $\omega$* is

$$F_\omega \coloneqq \big\{x \in H_1(M;\mathbb{R}) \;\big|\; \|x\|_{\mathrm{st}} = 1 \text{ and } \langle\omega, x\rangle = 1\big\} = H_\omega \cap \partial B\,,$$

and as in the preceeding subsection, we define *exposed points* and *exposed edges*.

As $F_\omega$ is the intersection of the closed and convex sets $B$ and $H_\omega$, the exposed face $F_\omega$ is closed and convex as well.

**Lemma 3.14.** *Let again $B \subset H_1(M;\mathbb{R})$ be the stable norm unit ball of a closed connected Riemannian manifold. If $\omega_0, \omega_1 \in H^1_{\mathrm{dR}}(M)$ with $\|\omega_0\|^*_{\mathrm{st}} = 1 = \|\omega_1\|^*_{\mathrm{st}}$ and*

*$F_{\omega_0} \cap F_{\omega_1} \neq \emptyset$, then for every $t \in (0,1)$ and for $\omega_t := (1-t)\omega_0 + t\omega_1 \in H^1_{\mathrm{dR}}(M)$ we have*

$$\|\omega_t\|^*_{\mathrm{st}} = 1 \quad \textit{and} \quad F_{\omega_t} = F_{\omega_0} \cap F_{\omega_1}\,.$$

*Proof.* We first show $\|\omega_t\|^*_{\mathrm{st}} = 1$: From the triangle inequality for $\|\bullet\|^*_{\mathrm{st}}$ we get $\|\omega_t\|^*_{\mathrm{st}} \leq 1$. Conversely, let $x \in F_{\omega_0} \cap F_{\omega_1}$. Then

$$\|\omega_t\|^*_{\mathrm{st}} \geq \frac{|\langle \omega_t, x\rangle|}{\|x\|_{\mathrm{st}}} = \frac{\bigl|(1-t)\langle \omega_0, x\rangle + t\langle \omega_1, x\rangle\bigr|}{1} = 1\,.$$

We clearly have $F_{\omega_0} \cap F_{\omega_1} \subset F_{\omega_t}$. Conversely, let $x \in F_{\omega_t}$. Then for $i \in \{0,1\}$ we have $|\langle \omega_i, x\rangle| \leq \|\omega_i\|^*_{\mathrm{st}} \cdot \|x\|_{\mathrm{st}} = 1$ by the defining equation (3.1). This implies

$$1 = \langle \omega, x\rangle = (1-t)\langle \omega_0, x\rangle + t\langle \omega_1, x\rangle \leq \bigl((1-t)\,\|\omega_0\|^*_{\mathrm{st}} + t\,\|\omega_1\|^*_{\mathrm{st}}\bigr) \cdot \|x\|_{\mathrm{st}} = 1\,.$$

Therefore, $\langle \omega_0, x\rangle = 1 = \langle \omega_1, x\rangle$, which shows that $x \in F_{\omega_0} \cap F_{\omega_1}$. □

## 4. Bounded difference of the minimal length and the stable norm

We show for general closed connected Riemannian manifolds that, on the integral part of homology, the difference between the minimal length and the stable norm is uniformly bounded. This generalizes a result of D. Burago, who considered the special case of tori [11].

**Theorem 4.1.** *Let $M$ be a closed connected Riemannian manifold. Then, there is a constant $D \in \mathbb{R}$ with*

$$\forall_{x \in H_1(M;\mathbb{Z})_{\mathbb{R}}} \quad \|x\|_{\mathrm{st}} \leq N(x) \leq \|x\|_{\mathrm{st}} + D\,.$$

*A concrete bound for $D$ is provided in Equation (4.4).*

The theorem can be considered as a continuous analogue of Remark 3.4. A modified version of this theorem is mentioned by Gromov [18, Sec. 4.21, Remarks$_+$(b), equation (△)]. Gromov knew that it follows from Burago's techniques [11], but did not provide details.

Burago's proof in the torus case uses the following splitting lemma going back to Nazarov, which is also the starting point for us.

**Lemma 4.2** (splitting lemma [11, Lemma 2])**.** *Let $V$ be an $\mathbb{R}$-vector space of dimension $b \in \mathbb{N}$ and let $\rho\colon [0,\ell] \to V$ be a continuous path. Then there exists a number $d \in \{0,\dots,\lceil b/2\rceil\}$ and pairwise disjoint closed subintervals $[\sigma_i,\tau_i]$ of $[0,\ell]$ for $i \in \{1,\dots,d\}$ with*

$$\sum_{i=1}^{d} \bigl(\rho(\tau_i) - \rho(\sigma_i)\bigr) = \frac{1}{2} \cdot \bigl(\rho(\ell) - \rho(0)\bigr)\,. \tag{4.1}$$

*We order these intervals in such a way that $0 \leq \sigma_1 < \tau_1 < \sigma_2 < \dots < \tau_d \leq \ell$ and call this sequence a* splitting partition *for $\rho$.*

The original proof by Nazarov was simplified by Burago [11], using ideas of Burago and Perelman. As our notation and statement slightly differs, we repeat Burago's argument for the sake of self-containedness.

*Proof.* Without loss of generality we assume $V = \mathbb{R}^b$. For a point $x = (x_1, \dots, x_{b+1}) \in S^b$ and $k \in \{0, 1, \dots, b+1\}$ we define

$$t_k(x) := \ell \cdot \sum_{i=1}^{k} {x_i}^2, \quad v(x) := \sum_{i=1}^{b+1} \operatorname{sign} x_i \cdot \big(\rho\big(t_i(x)\big) - \rho\big(t_{i-1}(x)\big)\big) .$$

Then $v\colon S^b \to \mathbb{R}^b$ is a continuous map with $v(-x) = -v(x)$ for all $x \in S^b$. By the Borsuk–Ulam theorem, there exists a $z = (z_1, \dots, z_{b+1}) \in S^b$ with $v(z) = 0$.

For $k \in \{1, \dots, b+1\}$, we consider the interval $I_k := [t_{k-1}(z), t_k(z)]$. Then, $I_k$ has positive length if and only if $z_k \neq 0$. Thus the sets

$$A_+ := \bigcup_{k \text{ with } z_k > 0} I_k \quad \text{and} \quad A_- := \bigcup_{k \text{ with } z_k < 0} I_k$$

satisfy

$$[0, \ell] = A_+ \cup A_-, \quad \mathring{A}_+ \cap \mathring{A}_- = \emptyset, \quad \overline{\mathring{A}_\pm} = A_\pm ,$$

and both $A_+$ and $A_-$ are finite unions of disjoint closed intervals of positive length. Then $A_+$ or $A_-$ is the union of $d \leq \lfloor (b+1)/2 \rfloor = \lceil b/2 \rceil$ such intervals, denoted as $[\sigma_i, \tau_i]$ with $0 \leq \sigma_1 < \tau_1 < \sigma_2 < \dots < \tau_d \leq \ell$. By construction, $v(z) = 0$ is equivalent to Equation (4.1). □

**Corollary 4.3.** *Let $M$ be a closed connected Riemannian manifold and $b := \dim_{\mathbb{R}} H_1(M; \mathbb{R})$. Let $\gamma$ be a continuous loop in $M$, viewed as a periodic continuous curve $\mathbb{R} \to M$ with period $\ell > 0$. Let $\widetilde{\gamma}$ be a lift of $\gamma$ to $\widetilde{M}$. Then, there is a real number $a \in [0, \ell)$, an integer $d \in \{0, 1, \dots, \lceil b/2 \rceil\}$, and real numbers $\sigma_i, \tau_i$ with $a = \sigma_1 < \tau_1 < \sigma_2 < \dots < \tau_d \leq a + \ell$ and*

$$\sum_{i=1}^{d} \big(J(\widetilde{\gamma}(\tau_i)) - J(\widetilde{\gamma}(\sigma_i))\big) = \frac{1}{2} \cdot \big(J(\widetilde{\gamma}(\ell)) - J(\widetilde{\gamma}(0))\big) = \frac{1}{2} \cdot h\big(\gamma|_{[0,\ell]}\big) .$$

*Proof.* We apply Lemma 4.2 to $J \circ \gamma$ and set $a := \sigma_1$. For the last equality, we use the relation between the geometric Hurewicz map and the Jacobi map (Remark 2.10). □

To prove Theorem 4.1, we introduce some geometric quantitites: Let $M$ be a closed connected Riemannian manifold and let $b := \dim_{\mathbb{R}} H_1(M; \mathbb{R})$. The operator norm $\|d_x J\|_{g,\mathrm{st}}$ of $d_x J\colon (T_x \widetilde{M}, g) \to (H_1(M; \mathbb{R}), \|\bullet\|_{\mathrm{st}})$ is continuous in $x \in \widetilde{M}$ (see Remark 2.10) and invariant under deck transformations of the universal covering $\widetilde{M} \to M$. Hence, there is a uniform upper bound and we may define

$$\mathcal{J} := \sup_{x \in \widetilde{M}} \|d_x J\|_{g,\mathrm{st}} < \infty . \tag{4.2}$$

**Lemma 4.4.** *In this situation, for every $C^\infty_{\mathrm{pw}}$-curve $\rho\colon [a, b] \to M$, we have*

$$\|h(\rho)\|_{\mathrm{st}} \leq \mathcal{J} \cdot \mathcal{L}(\rho) .$$

*Proof.* Using Remark 2.10, we obtain for a lift $\tilde\rho$ of $\rho$

$$\begin{aligned}\|h(\rho)\|_{\mathrm{st}} &= \big\|J\big(\tilde\rho(b)\big) - J\big(\tilde\rho(a)\big)\big\|_{\mathrm{st}} \\ &\le \int_a^b \Big\|\frac{d}{dt} J\big(\tilde\rho(t)\big)\Big\|_{\mathrm{st}}\, dt \\ &\le \int_a^b \big\|\mathrm{d}_{\tilde\rho(t)} J\big\|_{g,\mathrm{st}} \cdot \|\dot\rho(t)\|_g \, dt \\ &\le \mathcal{J}\cdot \mathcal{L}(\rho)\,. \end{aligned}$$ □

We now set

$$K := \sup\Big\{ N(y) \;\Big|\; y \in H_1(M;\mathbb{Z})_{\mathbb{R}},\ \|y\|_{\mathrm{st}} \le \mathcal{J}\cdot \frac{b+1}{2}\cdot \operatorname{diam} M \Big\}\,. \tag{4.3}$$

The following doubling property will provide the essential step for the proof of Theorem 4.1.

**Proposition 4.5** (Doubling property for the minimal length quasi-norm)**.** *The minimal length quasi-norm on a closed connected Riemannian manifold $M$ with $b = \dim_{\mathbb{R}} H_1(M;\mathbb{R})$ satisfies the doubling property from Definition 3.1 for the constant*

$$D := (b+5)\cdot \operatorname{diam} M + 2\cdot K\,. \tag{4.4}$$

*Here, $K$ is the constant from Equation (4.3).*

*Proof.* For all $x \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ and the constant $D$ from the claim, we have to show

$$2\cdot N(x) \le N(2\cdot x) + D\,.$$

For $x \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ we abbreviate $\ell := N(2\cdot x)$. We choose $\gamma\colon [0,\ell] \to M$ as a closed geodesic representing $2\cdot x$ with $\mathcal{L}(\gamma) = \ell = N(2\cdot x)$. Let $\widetilde\gamma$ be a lift of $\gamma$ to $\widetilde M$. To find "sufficiently short" representatives for $x$, we split $\widetilde\gamma$ via Corollary 4.3; by possibly reparametrizing $\gamma$ we may assume $a = 0$ in this corollary. For $i \in \{0,\dots,d-1\}$, we define $r_{2i} := \sigma_{i+1}$, $r_{2i+1} := \tau_{i+1}$, and $r_{2d} := \ell$. We thus have obtained a splitting partition $0 = r_0 < r_1 < \dots < r_k = \ell$ of $[0,\ell]$ with an even number $k := 2d \le b+1$.

Let $\vartheta_j := \gamma|_{[r_j,r_{j+1}]}$. The conclusion from Corollary 4.3 and Remark 2.10 give us that

$$\sum_{i=0}^{d-1} h(\vartheta_{2i}) = \sum_{i=0}^{d-1} h(\vartheta_{2i+1}) = \frac{1}{2}\cdot h(\gamma) = x\,.$$

We rearrange $\vartheta_0, \vartheta_2, \dots$ and $\vartheta_1, \vartheta_3, \dots$ into two loops: For $j \in \{0,\dots,k-1\}$, let $\varrho_j\colon [0,1] \to M$ be a smooth path from $\gamma(r_j)$ to $\gamma(r_{j+1})$ with $\mathcal{L}(\varrho_j) \le \operatorname{diam} M$. We consider the rearranged $C^\infty_{\mathrm{pw}}$-loops

$$\begin{aligned}\gamma_+ &:= \vartheta_0 * \rho_1 * \vartheta_2 * \dots * \vartheta_{k-2} * \rho_{k-1}\,,\\ \gamma_- &:= \rho_0 * \vartheta_1 * \rho_2 * \vartheta_3 * \dots * \vartheta_{k-1}\,.\end{aligned}$$

By construction, we have

$$\begin{aligned} y_+ \coloneqq h(\gamma_+) - \frac{1}{2}\cdot h(\gamma) &= \underbrace{\sum_{i=0}^{d-1} h(\vartheta_{2i})}_{=x} + \sum_{i=0}^{d-1} h(\rho_{2i+1}) - \frac{1}{2}\cdot 2\cdot x \\ &= \sum_{i=0}^{d-1} h(\rho_{2i+1})\,. \end{aligned}$$

As $\gamma_+$ is closed, we get $y_+ = h(\gamma_+) - x \in H_1(M;\mathbb{Z})_\mathbb{R}$, see Remark 2.7.

Our next goal will be to adapt $\gamma_+$ and $\gamma_-$ in such a way that the resulting loops are "sufficiently short" and represent $x$. We use Lemma 4.4 which tells us $\|h(\rho_j)\|_{\mathrm{st}} \le \mathcal{J}\cdot\mathcal{L}(\rho_j) \le \mathcal{J}\cdot \operatorname{diam} M$, and so

$$\|y_+\|_{\mathrm{st}} \;\le\; \sum_{i=0}^{d-1} \|h(\rho_{2i+1})\|_{\mathrm{st}} \;\le\; \mathcal{J}\cdot d\cdot \operatorname{diam} M\,.$$

In particular, $N(y_+) \le K$, by definition of $K$. Hence, we may find a loop $\mu_+$ with basepoint $\gamma_+(0)$ of length at most $K + 2\cdot\operatorname{diam} M$ satisfying $h(\mu_+) = -y_+$. In particular,

$$h(\gamma_+ * \mu_+) = y_+ + \frac{1}{2}\cdot h(\gamma) + h(\mu_+) = \frac{1}{2}\cdot h(\gamma) = x\,.$$

Similarly, we construct a loop $\mu_-$ from $\gamma_-$ with $\mathcal{L}(\mu_-) \le K + 2\cdot\operatorname{diam} M$ and $h(\gamma_- * \mu_-) = x$. For the corresponding lengths we calculate

$$\begin{aligned} \mathcal{L}(\gamma_+ * \mu_+) + \mathcal{L}(\gamma_- * \mu_-) &\le \sum_{i=0}^{k-1}\big(\mathcal{L}(\vartheta_i) + \mathcal{L}(\rho_i)\big) + \mathcal{L}(\mu_+) + \mathcal{L}(\mu_-) \\ &\le \ell + k\cdot\operatorname{diam} M + 2\cdot K + 4\cdot\operatorname{diam} M\,. \end{aligned}$$

Therefore,

$$2\cdot N(x) \le \ell + (k+4)\cdot\operatorname{diam} M + 2\cdot K = N(2\cdot x) + (b+5)\cdot\operatorname{diam} M + 2\cdot K\,.$$

This is the desired doubling property and the proposition follows. ☐

*Proof of Theorem 4.1.* In view of the Propositions 3.3 and 3.7 on homogenisations and the construction of the stable norm, it follows from the doubling property in Proposition 4.5 that Theorem 4.1 holds for the same constant $D$. ☐

Furthermore, we will make use of the following related estimate, complementing Lemma 4.4:

**Proposition 4.6.** *Let $M$ be a closed connected Riemannian manifold and let $h$ be a geometric Hurewicz map as constructed in Section 2.4. Then for all compact intervals $I \subset \mathbb{R}$ and all $C^\infty_{\mathrm{pw}}$-curves $\gamma\colon I \to M$, we have*

$$\mathcal{L}(\gamma) \ge \|h(\gamma)\|_{\mathrm{st}} - (\mathcal{J}+1)\cdot\operatorname{diam} M\,,$$

*where $\mathcal{J}$ is defined in Equation (4.2).*

*Proof.* We extend $\gamma$ to a $C^\infty_{\mathrm{pw}}$-loop $\gamma_0 \coloneqq \gamma * \varrho$, where $\varrho$ is a smooth path from the endpoint of $\gamma$ to the start point of $\gamma$ with $\mathcal{L}(\varrho) \le \operatorname{diam} M$. Then $\mathcal{L}(\gamma) = \mathcal{L}(\gamma_0) - \mathcal{L}(\varrho)$

and $h(\gamma_0) = h(\gamma) + h(\varrho)$. Moreover, $h(\gamma_0)$ is an integral class (Remark 2.7). We compute

$$\begin{aligned}
\mathcal{L}(\gamma) = \mathcal{L}(\gamma_0) - \mathcal{L}(\varrho) &\geq N\big(h(\gamma_0)\big) - \mathcal{L}(\varrho) \\
&\geq \|h(\gamma_0)\|_{\mathrm{st}} - \mathcal{L}(\varrho) \geq \|h(\gamma)\|_{\mathrm{st}} - \|h(\varrho)\|_{\mathrm{st}} - \mathcal{L}(\varrho) \\
&\geq \|h(\gamma)\|_{\mathrm{st}} - (\mathcal{J}+1)\cdot\mathcal{L}(\varrho) && \text{(by Lemma 4.4)} \\
&\geq \|h(\gamma)\|_{\mathrm{st}} - (\mathcal{J}+1)\cdot\operatorname{diam} M\,,
\end{aligned}$$

as claimed. □

We end the section with some observations that will help, in Subsection 5.3, to give a short proof of Proposition 5.5, but which are also of independent interest. The following propositions essentially generalize a corresponding result of Burago for tori [11, Theorem 1]. (We use the word “essentially” as Burago’s constant is controlled in terms of different geometric data.)

**Proposition 4.7** (Finite distance of Riemannian distance and stable norm)**.** *Let $(M, g)$ be a closed Riemannian manifold with universal covering $(\widetilde{M}, \widetilde{g})$, and let $\widetilde{d}$ be the induced distance function on $\widetilde{M}$. Let $J\colon \widetilde{M} \to H_1(M;\mathbb{R})$ be the Jacobi map (Definition 2.9).*

(1) *Then for all $p, q \in \widetilde{M}$ we have*

$$\widetilde{d}(p,q) \geq \|J(q) - J(p)\|_{\mathrm{st}} - (\mathcal{J}+1)\cdot\operatorname{diam} M\,,$$

*where $\mathcal{J}$ is defined in Equation (4.2).*

(2) *If $[\pi_1(M), \pi_1(M)]$ is finite, then there is a constant $C > 0$ auch that for all $p, q \in \widetilde{M}$ we have*

$$\widetilde{d}(p,q) \leq \|J(q) - J(p)\|_{\mathrm{st}} + C\,.$$

(3) *Let us now consider the covering $\widehat{M} \to M$ instead of the universal covering, where $\widehat{M}$ is either the covering $\widehat{M}^{\mathbb{Z}}$ or $\widehat{M}^{\mathbb{R}}$ (or any covering in between), defined in Section 2.5. The Jacobi map factors to $\widehat{J}\colon \widehat{M} \to H_1(M;\mathbb{R})$. Let $\hat{d}$ be the associated distance function on $\widehat{M}$. Then there is a constant $C > 0$ such that for all $\hat{p}, \hat{q} \in \widehat{M}$ we have*

$$\big|\hat{d}(\hat{p},\hat{q}) - \big\|\widehat{J}(\hat{q}) - \widehat{J}(\hat{p})\big\|_{\mathrm{st}}\big| \leq C\,.$$

Upper estimates for the constant $C$ in (2) and (3) can be read off in most cases from the proof below.

*Proof.* Ad (3): As $\widehat{M}$ is complete, the Hopf–Rinow theorem tells us that for all $\hat{p}, \hat{q} \in \widehat{M}$ there is a curve $\hat{\gamma}$ of length $\hat{d}(\hat{p},\hat{q})$ from $\hat{p}$ to $\hat{q}$ in $\widehat{M}$. Proposition 4.6 then yields

$$\hat{d}(\hat{p},\hat{q}) \geq \|h(\gamma)\|_{\mathrm{st}} - (\mathcal{J}+1)\cdot\operatorname{diam} M\,. \tag{4.5}$$

The converse inequality will first be shown for $\widehat{M} = \widehat{M}^{\mathbb{R}}$. Let $\pi\colon \widehat{M}^{\mathbb{R}} \to M$ the corresponding covering, whose deck transformation group will be identified canonically with $H_1(M;\mathbb{Z})_{\mathbb{R}}$. We may choose a path $\tau$ from $p \coloneqq \pi(\hat{q})$ to $q \coloneqq \pi(\hat{p})$, with $\mathcal{L}(\tau) \leq \operatorname{diam} M$. We choose $\hat{\gamma}$ as above and set $\gamma \coloneqq \pi \circ \hat{\gamma}$. The closed loop $\gamma * \tau$ represents a class in $H_1(M;\mathbb{Z})_{\mathbb{R}}$ and we choose a closed geodesic $\sigma$ in $M$ that minimizes length in this class. Thus $\mathcal{L}(\sigma) = N([\sigma]) = N([\gamma * \tau])$. We choose a path $\alpha$ from $\sigma(0) = \sigma(\ell)$ to $p$ in $M$ of length at most $\operatorname{diam} M$. Using Theorem 4.1, we obtain

$$\gamma' := \overline{\alpha} * \sigma * \alpha * \overline{\tau}$$

is a path from $p$ to $q$ such that $\gamma * \overline{\gamma'}$ represents 0 in $H_1(M;\mathbb{Z})_{\mathbb{R}}$. In other words, $\gamma'$ lifts to a path from $\hat{p}$ to $\hat{q}$. We obtain

$$\begin{aligned} \hat{d}(\hat{p},\hat{q}) &\leq \mathcal{L}(\gamma') \;=\; 2\cdot\mathcal{L}(\alpha) + \mathcal{L}(\sigma) + \mathcal{L}(\tau) \\ &\leq N([\gamma * \tau]) + 3\cdot \operatorname{diam} M \\ &\leq \big\|[\gamma*\tau]\big\|_{\mathrm{st}} + 3\cdot\operatorname{diam} M + D\,, \end{aligned}$$

where we used Theorem 4.1 in the last inequality for $D$ given by Equation (4.4). Let $\hat{\tau}$ be a lift of $\tau$ starting in $\hat{q}$; we denote its endpoint by $\hat{p}_1$. In particular, $\hat{\gamma} * \hat{\tau}$ is a path from $\hat{p}$ to $\hat{p}_1$, lifting $\gamma * \tau$. By construction and Remark 2.6 we have

$$[\gamma*\tau] = h(\gamma*\tau) = h(\gamma) + h(\tau) = \widehat{J}(\hat{q}) - \widehat{J}(\hat{p}) + h(\tau)\,.$$

From Lemma 4.4 we see that $\|h(\tau)\|_{\mathrm{st}} \leq \mathcal{J}\cdot\mathcal{L}(\tau) \leq \mathcal{J}\cdot\operatorname{diam} M$. Using the triangle inequality for $\|\bullet\|_{\mathrm{st}}$, we finally get

$$\begin{aligned} \hat{d}(\hat{p},\hat{q}) &\leq \big\|\widehat{J}(\hat{q}) - \widehat{J}(\hat{p})\big\|_{\mathrm{st}} + \|h(\tau)\|_{\mathrm{st}} + 3\cdot\operatorname{diam} M + D \\ &\leq \big\|\widehat{J}(\hat{q}) - \widehat{J}(\hat{p})\big\|_{\mathrm{st}} + (3+\mathcal{J})\cdot\operatorname{diam} M + D\,. \end{aligned}$$

This proves the converse inequality for $\widehat{M} = \widehat{M}^{\mathbb{R}}$.

As the torsion group of $H_1(M;\mathbb{Z})$ is finite, the covering $\widehat{M}^{\mathbb{Z}} \to \widehat{M}^{\mathbb{R}}$ is finite. We thus can apply Lemma A.4 for $Q_1 := \widehat{M}^{\mathbb{Z}}$ and $Q_2 := \widehat{M}^{\mathbb{R}}$, and this provides a constant $C > 0$ with

$$\forall_{\hat{p},\hat{q}\in\widehat{M}^{\mathbb{Z}}} \quad \hat{d}^{\mathbb{Z}}(\hat{p},\hat{q}) \;\leq\; \big\|\widehat{J}(\hat{q}) - \widehat{J}(\hat{p})\big\|_{\mathrm{st}} + \|h(\tau)\|_{\mathrm{st}} + (3+\mathcal{J})\cdot\operatorname{diam} M + D + C\,.$$

Ad (1): Let $p,q \in \widetilde{M}$ with images $\hat{p},\hat{q} \in \widehat{M}^{\mathbb{R}}$. Obviously, we have $\widetilde{d}(p,q) \geq \hat{d}(\hat{p},\hat{q})$. Thus, the claimed inequality directly follows from Equation (4.5).

Ad (2): The covering $\widetilde{M} \to \widehat{M}^{\mathbb{Z}}$ has deck transformation goup $[\pi_1(M),\pi_1(M)]$, and thus under the assumption of this item, this is a finite covering of bounded geometry. Lemma A.4 implies that there is a constant $C > 0$ such that for all $\hat{p},\hat{q} \in \widehat{M}^{\mathbb{Z}}$ with lifts $\widetilde{p},\widetilde{q} \in \widetilde{M}$ we have

$$\hat{d}(\hat{p},\hat{q}) \leq \widetilde{d}(\widetilde{p},\widetilde{q}) \leq \hat{d}(\hat{p},\hat{q}) + C\,.$$

The claim thus follows from Item (3). □

## 5. Homological asymptotes

We introduce homological asymptotes for curves and distinguish different types of minimal geodesics according to their homological asymptotes. Moreover, we collect various examples.

5.1. **Definition of homological asymptotes.** In most of the literature, the asymptotic direction of a minimal geodesic in homology was termed a "rotation vector". This notion is very intuitive in the case $b := \dim H_1(M;\mathbb{R}) = 2$ and, in particular, in the case of a 2-dimensional torus (Example 9.1). However, for $b > 2$ it seems more adequate to introduce the following asymptotes in $H_1(M;\mathbb{R})$. These are close to the definition of rotation vectors by Bangert [6] and the first author (Section 5.4).

**Definition 5.1** (terminal/initial/mixed asymptotes)**.** Let $M$ be a closed connected Riemannian manifold, let $h$ be a geometric Hurewicz map for $M$, and let $\gamma\colon \mathbb{R} \to M$ be a $C^\infty_{\mathrm{pw}}$-curve parametrized by arclength.

- We say that $v \in H_1(M;\mathbb{R})$ is *a terminal asymptote of* $\gamma$ if there exists a sequence $(t_i)_{i\in\mathbb{N}_0}$ in $\mathbb{R}$ with $\lim_{i\to\infty} t_i = \infty$ and
$$v = \lim_{i\to\infty} \frac{h\bigl(\gamma|_{[t_0,t_i]}\bigr)}{t_i - t_0}\,. \tag{5.1}$$
We write $A^+(\gamma)$ for the set of all terminal asymptotes of $\gamma$.
- We say that $v \in H_1(M;\mathbb{R})$ is *an initial asymptote of* $\gamma$ if there exists a sequence $(t_i)_{i\in\mathbb{N}_0}$ in $\mathbb{R}$ with $\lim_{i\to\infty} t_i = -\infty$ and
$$v = \lim_{i\to\infty} \frac{h\bigl(\gamma|_{[t_i,t_0]}\bigr)}{t_0 - t_i}\,.$$
We write $A^-(\gamma)$ for the set of all initial asymptotes of $\gamma$.
- We say that $v \in H_1(M;\mathbb{R})$ is *a mixed asymptote of* $\gamma$ if there are sequences $(s_i)_{i\in\mathbb{N}}$, $(t_i)_{i\in\mathbb{N}}$ in $\mathbb{R}$ with $\lim_{i\to\infty} s_i = -\infty$ and $\lim_{i\to\infty} t_i = \infty$ and
$$v = \lim_{i\to\infty} \frac{h\bigl(\gamma|_{[s_i,t_i]}\bigr)}{t_i - s_i}\,.$$
We write $A_{\mathrm{mix}}(\gamma)$ for the set of all mixed asymptotes of $\gamma$.

Moreover, we set $A^{+\cup-}(\gamma) := A^+(\gamma) \cup A^-(\gamma)$.

**Remark 5.2.** A straightforward triangle inequality estimate shows that terminal/initial asymptotes are independent of the start point $t_0$.

**Remark 5.3.** Let $M$ be a closed connected Riemannian manifold and let $\gamma, \sigma\colon \mathbb{R} \to M$ be minimal geodesics of $M$. If $\gamma$ and $\sigma$ are geometrically equivalent, say $\sigma = \gamma \circ \varphi$, then
$$\begin{aligned} &A^+(\gamma) = A^+(\sigma) \quad \text{and} \quad A^-(\gamma) = A^-(\sigma) && \text{if } \varphi' > 0\,,\\ \text{or}\quad &A^+(\gamma) = -A^-(\sigma) \quad \text{and} \quad A^-(\gamma) = -A^+(\sigma) && \text{if } \varphi' < 0\,. \end{aligned}$$

5.2. **Classical and elementary results about homological asymptotes.** The following lemma shows that the homological asymptotes for minimal geodesics lie in the (closed) unit ball of the stable norm. If $A$ and $B$ are a subsets of an affine space $V$, then we define their *convex join* as
$$\text{conv-join}(A,B) := \bigl\{ t\cdot x + (1-t)\cdot y \bigm| t \in [0,1] \text{ and } x \in A \text{ and } y \in B \bigr\}\,,$$
which is a subset of the convex hull $\mathrm{conv}(A \cup B)$.

**Lemma 5.4.** *Let $M$ be a closed connected Riemannian manifold and let $\gamma\colon \mathbb{R} \to M$ be a $C^\infty_{\mathrm{pw}}$-curve parametrized by arclength.*

(1) *Let $(t_i)_{i\in\mathbb{N}_0}$ be a sequence in $\mathbb{R}$ with $\lim_{i\to\infty} t_i = \infty$. Then*
$$\limsup_{i\to\infty} \frac{\bigl\| h(\gamma|_{[t_0,t_i]})\bigr\|_{\mathrm{st}}}{t_i - t_0} \leq 1\,.$$
(2) *Thus, $A^{+\cup-}(\gamma)$ lies in the closed unit ball of $(H_1(M;\mathbb{R}), \|\cdot\|_{\mathrm{st}})$.*
(3) *If $(t_i)_{i\in\mathbb{N}_0}$ is a sequence in $\mathbb{R}$ with $\lim_{i\to\infty} t_i = \infty$, then there exists a subsequence $(t'_i)_{i\in\mathbb{N}}$ of $(t_i)_{i\in\mathbb{N}}$ such that $\lim_{i\to\infty} h\bigl(\gamma|_{[t_0,t'_i]}\bigr)/(t'_i - t_0)$ exists.*

(4) *In particular, $A^+(\gamma) \neq \emptyset$ and $A^-(\gamma) \neq \emptyset$.*
(5) *The sets $A^+(\gamma)$ and $A^-(\gamma)$ are closed.*
(6) *The sets $A^+(\gamma)$ and $A^-(\gamma)$ are connected.*
(7) *We have*

$$A^{+\cup-}(\gamma) \subset A_{\mathrm{mix}}(\gamma) \subset \text{conv-join}\big(A^+(\gamma), A^-(\gamma)\big)\,.$$

(8) *The set $A_{\mathrm{mix}}(\gamma)$ lies in the closed unit ball of $(H_1(M;\mathbb{R}), \|\bullet\|_{\mathrm{st}})$ and is closed and conected.*

*Proof.* Ad (1): As $\gamma$ is a minimal geodesic, we have $\mathcal{L}(\gamma|_{[t_0,t_i]}) = t_i - t_0$ for all $i \in \mathbb{N}$. Using the estimate between lengths and the stable norm (Proposition 4.6), we obtain a constant $c \in \mathbb{R}_{\geq 0}$ such that for all $i \in \mathbb{N}$:

$$\frac{\big\|h(\gamma|_{[t_0,t_i]})\big\|_{\mathrm{st}}}{t_i - t_0} \leq \frac{\mathcal{L}(\gamma|_{[t_0,t_i]}) + c}{t_i - t_0} = \frac{t_i - t_0 + c}{t_i - t_0}$$

Because of $\lim_{i\to\infty} t_i = \infty$, the claim follows.

Ad (2): We first prove the statement for $A^+(\gamma)$. Item (1) can be restated as

$$\limsup_{i\to\infty} \left\| \frac{h(\gamma|_{[t_0,t_i]})}{t_i - t_0} \right\|_{\mathrm{st}} \leq 1\,. \tag{5.2}$$

Thus if $v \in A^+(\gamma)$ is the limit for such a sequence $(t_i)_i$, then $\|v\|_{\mathrm{st}} \leq 1$. The proof for $A^-(\gamma)$ is analogous.

Ad (3): Because of Estimate (5.2) we see that $\big(h(\gamma|_{[t_0,t_i]})/(t_i - t_0)\big)_{i\in\mathbb{N}}$ lies in a stable norm ball of finite radius. Therefore, compactness of this ball allows us to select a subsequence that converges.

Ad (4): For $A^+(\gamma)$, we apply Item (3) to the sequence $(i)_{i\in\mathbb{N}}$. Reversing the orientation, we also get $A^-(\gamma) \neq \emptyset$.

Ad (5): We prove the statement for $A^+(\gamma)$. Let $v$ be in the closure of $A^+(\gamma)$, i.e., we have a sequence $(v_i)_{i\in\mathbb{N}}$ in $A^+(\gamma)$ with $v_i \to v$ for $i \to \infty$. For each $i$, we choose a $t_i$ with

$$\left\| v_i - \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i} \right\|_{\mathrm{st}} \leq \frac{1}{i}\,.$$

One can choose the $t_i$ such that $t_i \nearrow \infty$. Then for this sequence $(t_i)_i$ and $v$ we have the relation in Equation (5.1); thus, $v \in A^+(\gamma)$. The proof of the closedness of $A^-(\gamma)$ is analogous.

Ad (6): Assume for a contradiction that $A^+(\gamma)$ were not connected, i. e., there exists open subsets $U_0$ and $V_0$ of $H_1(M;\mathbb{R})$ with

$$A^+(\gamma) \cap U_0 \cap V_0 = \emptyset,\quad A^+(\gamma) \subset U_0 \cup V_0 \quad\text{and}\quad A^+(\gamma) \cap U_0 \neq \emptyset \neq A^+(\gamma) \cap V_0\,. \tag{5.3}$$

Then $A^+(\gamma) \cap U_0 = A^+(\gamma) \setminus V_0$ is closed and thus compact. Similarly $A^+(\gamma) \cap V_0$ is compact. We thus can replace $U_0$ and $V_0$ in (5.3) by open subsets $U$ and $V$ with $U \cap V = \emptyset$. We choose $u \in A^+(\gamma) \cap U$ and $v \in A^+(\gamma) \cap V$. There are sequences $(s_i)_{i\in\mathbb{N}}$ and $(t_i)_{i\in\mathbb{N}}$ in $\mathbb{R}_{>0}$ with $\lim_{i\to\infty} s_i = \infty = \lim_{i\to\infty} t_i$ such that

$$u = \lim_{i\to\infty} \frac{h\big(\gamma|_{[0,s_i]}\big)}{s_i} \quad\text{and}\quad v = \lim_{i\to\infty} \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i}\,.$$

Passing to subsequences we can achieve $s_1 < t_1 < s_2 < t_2 < s_2 \ldots$ and

$$\lim_{i\to\infty} \frac{h\big(\gamma|_{[0,s_i]}\big)}{s_i} \in U \quad\text{and}\quad \lim_{i\to\infty} \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i} \in V\,.$$

By continuity, for each $i \in \mathbb{N}$, we can choose an $r_i \in (s_i, t_i)$ with

$$w_i := \frac{h\big(\gamma|_{[0,r_i]}\big)}{r_i} \notin U \cup V\,.$$

Moreover, by Lemma 4.4, we have $w_i \in \overline{B}_{\mathcal{J}}(0)$, where $\mathcal{J}$ is defined as in Equation (4.2). Thus, after passing to further subsequences we obtain that $w_\infty := \lim_{i\to\infty} w_i$ exists and lies in $\overline{B}_{\mathcal{J}}(0) \setminus (U \cap V)$. (Alternatively, we could use Item (3) here again.) By defintion of $A^+(\gamma)$, we also have $w_\infty \in A^+(\gamma)$, which contradicts the assumption that $A^+(\gamma) \subset U \cup V$.

By reversing orientations, we also see that $A^-(\gamma)$ is connected.

Ad (7): We first show the inclusion $A^{+\cup-}(\gamma) \subset A_{\mathrm{mix}}(\gamma)$. For a given $v \in A^+(\gamma)$ we choose a sequence $t_i \nearrow \infty$ with $h\big(\gamma|_{[0,t_i]}\big)/t_i \to v$. After passing to a subsequence, we may assume $t_i \geq i^2$ for all $i \in \mathbb{N}$. With Lemma 4.4 we see that $\big\|h\big(\gamma|_{[a,b]}\big)\big\|_{\mathrm{st}} \leq \mathcal{J} \cdot (b-a)$. We get

$$\begin{aligned}
\left\| \frac{h\big(\gamma|_{[-i,t_i]}\big)}{t_i+i} - \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i} \right\|_{\mathrm{st}} &\leq \left\| \left(\frac{t_i}{t_i+i} - 1\right) \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i} \right\|_{\mathrm{st}} + \left\| \frac{h\big(\gamma|_{[-i,0]}\big)}{t_i+i} \right\|_{\mathrm{st}} \\
&\leq \left| \frac{i}{t_i+i} \right| \cdot \frac{\big\|h\big(\gamma|_{[0,t_i]}\big)\big\|_{\mathrm{st}}}{t_i} + \frac{\mathcal{J}\cdot i}{|t_i+i|} \\
&\leq \frac{i}{i^2+i} \cdot \mathcal{J} + \frac{\mathcal{J}\cdot i}{i^2+i} \leq \frac{2\mathcal{J}}{i}
\end{aligned}$$

and thus

$$v = \lim_{i\to\infty} \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i} = \lim_{i\to\infty} \frac{h\big(\gamma|_{[-i,t_i]}\big)}{t_i+i} \in A_{\mathrm{mix}}(\gamma)\,.$$

So we have $A^+(\gamma) \in A_{\mathrm{mix}}(\gamma)$. The proof for the inclusion $A^-(\gamma) \subset A_{\mathrm{mix}}(\gamma)$ is obtained by reversing the orientation. In total, we get $A^{+\cup-}(\gamma) \subset A_{\mathrm{mix}}(\gamma)$.

Now, we will prove the inclusion $A_{\mathrm{mix}}(\gamma) \subset \text{conv-join}\big(A^+(\gamma), A^-(\gamma)\big)$. Suppose that $v \in A_{\mathrm{mix}}(\gamma)$ may be obtained as

$$v = \lim_{i\to\infty} \frac{h\big(\gamma|_{[s_i,t_i]}\big)}{t_i - s_i}\,.$$

with $\lim_{i\to\infty} s_i = -\infty$ and $\lim_{i\to\infty} t_i = \infty$. By Item (3), after passing to subsequences, the individual limits

$$\tau := \lim_{i\to\infty} \frac{t_i}{t_i - s_i}, \quad v_+ := \lim_{i\to\infty} \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i}, \quad \text{and} \quad v_- := \lim_{i\to\infty} \frac{h\big(\gamma|_{[s_i,0]}\big)}{-s_i}$$

exist and we have $\tau \in [0,1]$, $v_+ \in A^+(\gamma)$, and $v_- \in A^-(\gamma)$. We calculate

$$\begin{aligned}
\frac{h\big(\gamma|_{[s_i,t_i]}\big)}{t_i - s_i} &= \frac{t_i}{t_i - s_i} \frac{h\big(\gamma|_{[0,t_i]}\big)}{t_i} + \frac{-s_i}{t_i - s_i} \frac{h\big(\gamma|_{[s_i,0]}\big)}{-s_i} \\
&\xrightarrow{i\to\infty} \tau \cdot v_+ + (1-\tau)\cdot v_- \in \text{conv-join}\big(A^+(\gamma), A^-(\gamma)\big)\,.
\end{aligned}$$

Ad (8): It follows from Items (2) and (7), and from the convexity of norm balls, that $A_{\mathrm{mix}}(\gamma)$ is contained in the stable norm unit ball.

To prove closedness of $A_{\mathrm{mix}}(\gamma)$ we proceed as in Part (5) with $A^+$ replaced by $A_{\mathrm{mix}}$, and with $h(\gamma|_{[0,t_i]})/t_i$ replaced by $h(\gamma|_{[s_i,t_i]})/(t_i - s_i)$. Again the choices can be done such that $t_i \nearrow \infty$ and $s_i \searrow -\infty$. As before, this implies $v \in A_{\mathrm{mix}}(\gamma)$.

For connectedness, one can use similar arguments as in the proof of Item (6). □

5.3. **Asymptotes of homologically minimal geodesics.** Asymptotes of $\mathbb{Z}$-homologically minimal geodesics (Definition 2.12) lie in the stable norm unit sphere:

**Proposition 5.5.** *Let $M$ be a closed connected Riemannian manifold and let $\gamma\colon \mathbb{R} \to M$ be a $\mathbb{Z}$-homologically minimal geodesic of $M$. Then $A^+(\gamma)$, $A^-(\gamma)$, and $A_{\mathrm{mix}}(\gamma)$ are subsets of the stable norm unit sphere*

$$\left\{x \in H_1(M;\mathbb{R}) \;\middle|\; \|x\|_{\mathrm{st}} = 1\right\}.$$

As $\mathbb{R}$-homological minimality implies $\mathbb{Z}$-homological minimality, the proposition also holds for all $\mathbb{R}$-homologically minimal geodesics.

*Proof.* It suffices to prove the statement for $A_{\mathrm{mix}}(\gamma)$. Let $v \in A_{\mathrm{mix}}(\gamma)$. As a $\mathbb{Z}$-homologically minimal geodesic is minimal, we know that $\|v\|_{\mathrm{st}} \le 1$ from Lemma 5.4 (8).

We prove $\|v\|_{\mathrm{st}} \ge 1$: For $L > 0$, we choose $s, t \in \mathbb{R}$ with $s + L < t$ such that

$$\left\|v - \frac{h\big(\gamma|_{[s,t]}\big)}{t-s}\right\|_{\mathrm{st}} < \varepsilon \coloneqq \frac{1}{L}.$$

Let $\hat\gamma$ be a lift of $\gamma$ to $\widehat{M}^{\mathbb{Z}}$. We apply Proposition 4.7 (3) for $\widehat{M} = \widehat{M}^{\mathbb{Z}}$, using the definitions of $\widehat{J}$, $\hat d$ and $C$ from this proposition, and we obtain:

$$\begin{aligned}
(\|v\|_{\mathrm{st}} + \varepsilon)\cdot(t-s) &\ge \left\|h\big(\gamma|_{[s,t]}\big)\right\|_{\mathrm{st}}\\
&\ge \left\|\widehat{J}\big(\hat\gamma(t)\big) - \widehat{J}\big(\hat\gamma(s)\big)\right\|_{\mathrm{st}}\\
&\ge \hat d\big(\hat\gamma(t), \hat\gamma(s)\big) - C\\
&\ge (t-s) - C.
\end{aligned}$$

This yields

$$\|v\|_{\mathrm{st}} \ge 1 - \frac{C}{t-s} - \varepsilon \ge 1 - \frac{C}{L} - \varepsilon$$

and in the limit $L = \varepsilon^{-1} \to \infty$ we finally obtain $\|v\|_{\mathrm{st}} \ge 1$. □

5.4. **Comparison to Bangert's rotation vectors.** Initial and terminal asymptotes are related to Bangert's "rotation vectors" [6, 1]. For historical clarity, let us explain this, although it is not required for the article's main results.

For a curve $\gamma\colon [a,b] \to M$, Bangert defines the *rotation vector* as

$$R(\gamma) \coloneqq \frac{h(\gamma)}{\|h(\gamma)\|_{\mathrm{st}}}.$$

Now let $\gamma\colon \mathbb{R} \to M$ be a minimal geodesic. In Bangert's work [6], the accumulation points of $R\big(\gamma|_{[s,t]}\big)$ with $t - s \to \infty$ play an important role. If additionally $\gamma$ is $\mathbb{Z}$-homologically minimal, then he shows that these accumulation points are in the intersection of the stable norm unit sphere and a supporting hyperplane which implies a statement similar to our Proposition 5.5 [6, Theorem 3.2]. More precisely:

**Theorem 5.6** (Bangert [6, Theorem 3.2])**.** *Let $M$ be a closed connected Riemannian manifold and let $B$ be the stable norm unit ball in $H_1(M;\mathbb{R})$. Let $\gamma$ be a $\mathbb{Z}$-homologically minimal geodesic. Then there exists a supporting hyperplane $H$ to $B$ with the following property: for every neighborhood $U$ of $H \cap B$ there exists $c_U > 0$ such that $R\big(\gamma|_{[s,t]}\big) \in U$ whenever $t - s > c_U$.*

Conversely, Bangert proves that for every supporting hyperplane a $\mathbb{Z}$-homologically minimal geodesic exists such that the accumulation points of the rotation vectors lie in the intersection of that hyperplane with the stable norm unit sphere [6, Theorem 4.4].

Rotation vectors and asymptotes are related as follows:

**Proposition 5.7.** *Let $M$ be a closed connected Riemannian manifold and let $\gamma\colon \mathbb{R} \to M$ be a $\mathbb{Z}$-homologically minimal geodesic. Then there are constants $C_0$, $c_0 > 0$ such that*

$$\forall_{s,t\in\mathbb{R}} \quad t-s \geq c_0 \Longrightarrow \left\| R\bigl(\gamma|_{[s,t]}\bigr) - \frac{h(\gamma|_{[s,t]})}{t-s} \right\|_{\mathrm{st}} \leq \frac{C_0}{t-s}\,.$$

*Proof.* Let $\hat\gamma$ be a lift of $\gamma$ to the covering $\widehat{M}^{\mathbb{Z}}$. By $\mathbb{Z}$-homological minimality, $\hat\gamma$ is a geodesic line, i. e., for all $s<t$ we have $\hat d\bigl(\hat\gamma(t),\hat\gamma(s)\bigr) = t-s$.

We calculate for the constant $C$ provided by Proposition 4.7 (3) that

$$\begin{aligned} \Bigl| \bigl\| h\bigl(\gamma|_{[s,t]}\bigr)\bigr\|_{\mathrm{st}} - (t-s)\Bigr| &= \Bigl| \bigl\| \widehat J\bigl(\hat\gamma(t)\bigr) - \widehat J\bigl(\hat\gamma(s)\bigr)\bigr\|_{\mathrm{st}} - (t-s)\Bigr| \\ &\leq \underbrace{\Bigl| \hat d\bigl(\hat\gamma(t),\hat\gamma(s)\bigr) - (t-s)\Bigr|}_{=0} + C \;=\; C\,. \end{aligned}$$

Moreover, we have

$$\begin{aligned} \left\| R\bigl(\gamma|_{[s,t]}\bigr) - \frac{h(\gamma|_{[s,t]})}{t-s} \right\|_{\mathrm{st}} &\leq \left| 1 - \frac{\bigl\|h(\gamma|_{[s,t]})\bigr\|_{\mathrm{st}}}{t-s} \right| \cdot \bigl\| R\bigl(\gamma|_{[s,t]}\bigr)\bigr\|_{\mathrm{st}} \\ &\overset{(*)}{\leq} \frac{C\cdot(1+\varepsilon)}{t-s}\,; \end{aligned}$$

for the inequality $(*)$ we used that Bangert's Theorem 5.6 implies $\bigl\|R(\gamma|_{[s,t]})\bigr\|_{\mathrm{st}} \leq 1+\varepsilon$ for some $\varepsilon$ that goes to 0 uniformly in $t-s\to\infty$. (Alternatively, we could use Proposition 5.5 for that purpose). The statement follows for $C_0 \coloneqq C(1+\varepsilon)$. □

**Corollary 5.8.** *We assume the conditions of Proposition 5.7. Then*

$$A_{\mathrm{mix}}(\gamma) = \Bigl\{ \lim_{i\to\infty} R\bigl(\gamma|_{[s_i,t_i]}\bigr) \Bigm| t_i \nearrow +\infty \text{ and } s_i \searrow -\infty \text{ and the limit exists} \Bigr\}\,.$$

*Proof.* This is immediate from Proposition 5.7. □

We obtain the following improvement of Proposition 5.5:

**Corollary 5.9.** *Let $M$ be a closed connected Riemannian manifold and let $\gamma\colon\mathbb{R}\to M$ be a $\mathbb{Z}$-homologically minimal geodesic. Let $B$ be the (closed) stable norm unit ball. Then there is a supporting hyperplane $H$ for $B$ with*

$$\operatorname{conv} A^{+\cup-}(\gamma) \subset \bigl\{ x \in H_1(M;\mathbb{R}) \bigm| \|x\|_{\mathrm{st}} = 1 \bigr\} \cap H\,.$$

*Proof.* As the right-hand side is convex, it suffices to prove the statement for $A^+(\gamma)$ and $A^-(\gamma)$ instead of $\operatorname{conv} A^{+\cup-}(\gamma)$. We already know from Proposition 5.5 that $A^\pm(\gamma) \subset \bigl\{x \in H_1(M;\mathbb{R}) \bigm| \|x\|_{\mathrm{st}} = 1\bigr\}$. From Theorem 5.6 and Proposition 5.7, we get a supporting hyperplane $H$ such that $A^\pm(\gamma)\subset H$; the statement is thus shown. □

**Remark 5.10.** For general minimal geodesics (i. e., without the assumption "$\mathbb{Z}$-homologically"), the relationship of rotation vectors and asymptotes is not so immediate and they are out of the main focus of Bangert's work [6]. We state some relations without proof: If a vector $v \in H_1(M;\mathbb{R})$ is an accumulation point for $R\big(\gamma|_{[s_i,t_i]}\big)$ for some $\mathbb{R}$-valued sequences $(s_i)_{i\in\mathbb{N}}$ and $(t_i)_{i\in\mathbb{N}}$ with $t_i \nearrow +\infty$ and $s_i \searrow -\infty$, then there is a $w \in A_{\mathrm{mix}}(\gamma)$ auch that

$$\|w\|_{\mathrm{st}} \cdot v = w\,. \tag{5.4}$$

Note that we did not exclude the case $w = 0$, which arises in examples (Example 9.2). Conversely, for $w \in A_{\mathrm{mix}}(\gamma) \setminus \{0\}$, $v \coloneqq w/\|w\|_{\mathrm{st}}$ is an accumulation point for $R\big(\gamma|_{[s_i,t_i]}\big)$ with $t_i \nearrow +\infty$ and $s_i \searrow -\infty$.

5.5. **Homologically homoclinic and heteroclinic minimal geodesics.** The definitions in this section apply in any dimension, but they are motivated by Hedlund examples, which exist for $\dim M \geq 3$. Several notions introduced in this subsection come from dynamics, as – besides the question of existence of minimal geodesics – the dynamical properties of the geodesic flow close to minimal geodesics are of interest.

Hedlund examples [19, Section 9][6, Section 5][1, 2] have a finite number of simple closed geodesics; if parametrized by arclength as a periodic curve $\tau_i \colon \mathbb{R} \to M$, these simple closed geodesiscs are also minimal geodesics, see Section 9.2. In the present article, such curves will be called *minimal closed geodesics*. Let us mention that the notion "minimizing closed geodesics" is also used in the literature. If $(M, g)$ is a Hedlund example, one has a finite number $\ell$ of minimal closed geodesics, thus we assume $i \in \{1, 2, \cdots, \ell\}$.

**Definition 5.11.** Let $(M, g)$ be a closed Riemannian manifold with closed minimal geodesics $\tau_1, \dots, \tau_\ell$.

A minimal geodesic $\gamma$, parametrized by arclength, is called *asymptotic* to $\tau_i$ at $\pm\infty$ if there is a sign $\varepsilon \in \{-1, 1\}$ and an $\alpha \in \mathbb{R}$ such that

$$\lim_{t\to\pm\infty} d\big(\gamma(t), \tau_i\big(\varepsilon \cdot (t-\alpha)\big)\big) = 0\,.$$

If $\gamma$ is a minimal geodesic, following Poincaré we say that $\gamma$ is

- *homoclinic* if there is an $i \in \{1, \dots, \ell\}$ such that $\gamma$ is asymptotic to $\tau_i$ at $+\infty$ and $-\infty$. This includes the cases $\gamma = \tau_i$ and $\gamma = \overline{\tau}_i$.
- *heteroclinic* if there are $i_+, i_- \in \{1, \dots, \ell\}$ with $i_+ \neq i_-$ such that $\gamma$ is asymptotic to $\tau_{i_\pm}$ at $\pm\infty$.

If the fundamental group $\pi_1(M)$ of a closed connected manifold $M$ is virtually abelian (or virtually nilpotent of "bounded minimal generation" [2, Definition 7.1][1, Definition IV.1.19]), then for a suitable Hedlund metric on $M$, every minimal geodesic is either homoclinic or heteroclinic [1, Hilfsatz IV.1.9 and Korollar IV.1.21].

**Historical comment 5.12.** The existence of homoclinic and heteroclinic minimal geodesics was also studied by Bolotin and Rabinowitz, in the special case $M = T^n$, under some weak assumptions. We will explain this in Subsection 9.4 in more details.

On general Riemannian manifolds, this motivates the introduction of similar terminology, which will be useful to distinguish different classes of minimal geodesics.

All these notions only depend on the geometric equivalence class of the minimal geodesic.

**Definition 5.13** (types of minimal geodesics)**.** A minimal geodesic $\gamma$ on a closed connected Riemannian manifold $M$ is called

- *homologically homoclinic* if there is a $v \in H_1(M;\mathbb{R})$ with
$$A^+(\gamma) = \{v\} = A^-(\gamma)\,.$$
- *homologically heteroclinic* if there are $v, w \in H_1(M;\mathbb{R})$ with $v \neq w$ and
$$A^+(\gamma) = \{v\} \quad \text{and} \quad A^-(\gamma) = \{w\}\,.$$
- *homologically diverging* if $A^+(\gamma)$ or $A^-(\gamma)$ (or both of them) contain more than one point.
- *homologically semi-converging* if $A^+(\gamma)$ or $A^-(\gamma)$ (or both of them) contain precisely one point.
- *homologically exposed* if there are exposed points $x, y$ in the stable norm unit ball $B$ with $A^+(\gamma) = \{x\}$ and $A^-(\gamma) = \{y\}$. We allow both the homoclinic ($x = y$) and the hereroclinic ($x \neq y$) case.
- *homologically semi-exposed* if there is an exposed point $x$ of $B$ with $A^+(\gamma) = \{x\}$ or $A^-(\gamma) = \{x\}$.
- *homologically non-homoclinic* if it is not homologically homoclinic. By definition, this is equivalent to being either homologically heteroclinic or homologically diverging.

All homoclinic/heteroclinic minimal geodescis in the sense of Definition 5.11 are homologically homoclinic/heteroclinic, respectively, provided that the homology classes $[\tau_1], \dots, [\tau_\ell] \in H_1(M, \mathbb{R})$ are pairwise different. By definition, every minimal geodesic is either homologically homoclinic, homologically heteroclinic or homologically diverging. Furthermore, “homologically semi-exposed” implies “homologically semi-converging”.

**Example 5.14.** Let the torus $T^k = \mathbb{R}^k/\mathbb{Z}^k$ carry a flat metric and let $N$ be a closed connected Riemannian manifold with finite $\pi_1(N)$. We equip $M := T^k \times N$ with the product metric. Then a curve $\gamma = (\gamma_T, \gamma_N)$ is a geodesic if and only if both $\gamma_T$ and $\gamma_N$ are geodesics, not necessarily parametrized by arclength. Further, $\gamma$ is a minimal geodesic if and only if $\gamma_N$ is constant and $\gamma_T$ a minimal geodesic. The latter condition is equivalent to saying that after identifying $\widetilde{T^k}$ with Euclidean $\mathbb{R}^k$, the component $\gamma_T$ is an affine line. In particular, all minimal geodesics on $M$ are $\mathbb{R}$-homologically minimal, homologically homoclinic, homologically (semi-)converging, and homologically (semi-)exposed.

## 6. Finding new minimal geodesics

We explain how points on the exposed edges of the stable norm unit ball lead to minimal geodesics and how their terminal/initial asymptotes are controlled by the corresponding exposed edge. In particular, we will see how these minimal geodesics can be distinguished via their terminal/initial asymptotes.

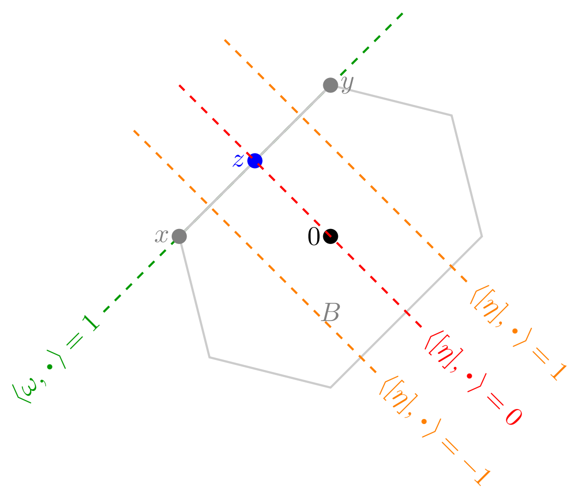


FIGURE 1. The situation in the construction of Section 6.2

6.1. **Setup for the construction.** Let $M$ be a closed connected Riemannian manifold, let $b := \dim_{\mathbb{R}} H_1(M;\mathbb{R})$, and let $B \subset H_1(M;\mathbb{R})$ be the stable norm unit ball, i.e., the closed unit ball of $\big(H_1(M;\mathbb{R}), \|\bullet\|_{\mathrm{st}}\big)$. The boundary of $B$ is the stable norm unit sphere $\partial B = \big\{x \in H_1(M;\mathbb{R}) \bigm| \|x\|_{\mathrm{st}} = 1\big\}$. An *exposed edge* of $B$ is a subset $E \subset \partial B$ that is the convex hull of two different points $x, y \in \partial B$ such that there is an $\omega \in H^1_{\mathrm{dR}}(M)$ with $\|\omega\|^*_{\mathrm{st}} = 1$ with $E = F_\omega$, where $F_\omega$ is defined in Definition 3.13. The second condition can be equally expressed by saying that there is a hyperplane $H$ of $H_1(M;\mathbb{R})$ supporting $B$ and satisfying $E = H \cap B = H \cap \partial B$. We will write $[x, y]$ for the convex hull of $x$ and $y$, viewed as a 1-dimensional submanifold with boundary with an orientation from $x$ any $y$. Hence $[x, y]$ and $[y, x]$ are equal as submanifolds, but with opposite orientations. Moreover, we write $(x, y) := [x, y] \setminus \{x, y\}$ for the relative interior of $[x, y]$. If $\dim_{\mathbb{R}} H_1(M;\mathbb{R}) \leq 1$, then $B$ does not have any exposed edges.

Moreover, we choose a $\mathbb{Z}$-module basis $(\beta_1, \ldots, \beta_b)$ of $H_1(M;\mathbb{Z})_{\mathbb{R}}$ and closed 1-forms $\alpha^1, \ldots, \alpha^b$ with $\langle [\alpha^i], \beta_j \rangle = \delta_{ij}$ as in Section 2.4; in particular, we obtain an associated geometric Hurewicz map $h$.

6.2. **Construction of minimal geodesics from exposed edges of the stable norm unit ball.** Let $[x, y] \subset \partial B$ be an exposed edge of the stable norm unit ball $B$ (Figure 1) and choose $z \in (x, y)$. By definition of an exposed edge – Definition 3.13 – there is an $\omega \in H^1_{\mathrm{dR}}(M)$ with $\|\omega\|^*_{\mathrm{st}} = 1$ and

$$\langle \omega, z \rangle = 1 = \|z\|_{\mathrm{st}} \quad \text{and} \quad [x, y] = F_\omega\,.$$

In addition, we choose a closed smooth 1-form $\eta \in \Omega^1(M)$ that is a linear combination of $\alpha^1, \ldots, \alpha^b$ and that satisfies

$$\langle [\eta], z \rangle = 0 \quad \text{and} \quad \langle [\eta], x \rangle < 0 < \langle [\eta], y \rangle\,.$$

As starting point for our construction, we pick a controlled approximation of $z$:

**Lemma 6.1.** *In this situation, there exists a sequence $(z_i)_{i\in\mathbb{N}}$ in $H_1(M;\mathbb{Z})_\mathbb{R} \setminus \{0\}$ and a constant $a \in \mathbb{R}_{\geq 0}$ such that*

$$\lim_{i\to\infty} \frac{z_i}{\|z_i\|_{\mathrm{st}}} = z, \quad \forall_{i\in\mathbb{N}} \quad \langle \omega, z_i\rangle \geq N(z_i) - a, \quad \textit{and} \quad \lim_{i\to\infty} \|z_i\|_{\mathrm{st}} = \infty\,.$$

The proof even gives the additional information that $\{\|z_i - i\cdot z\|_{\mathrm{st}} \mid i \in \mathbb{N}\}$ is bounded, but this will not be needed later.

*Proof.* Because $H_1(M;\mathbb{Z})_\mathbb{R}$ is a cocompact lattice in $H_1(M;\mathbb{R})$, there exists an $r \in \mathbb{R}_{>0}$ with

$$\forall_{v\in H_1(M;\mathbb{R})} \quad \exists_{v'\in H_1(M;\mathbb{Z})_\mathbb{R}} \quad \|v - v'\|_{\mathrm{st}} \leq r\,.$$

For each $i \in \mathbb{N}$, we choose $z_i \in H_1(M;\mathbb{Z})_\mathbb{R}$ with $\|i\cdot z - z_i\|_{\mathrm{st}} \leq r$.

By construction,

$$\begin{aligned}\|z_i\|_{\mathrm{st}} &\leq \|i\cdot z\|_{\mathrm{st}} + \|z_i - i\cdot z\|_{\mathrm{st}} \leq i + r\,,\\ \|z_i\|_{\mathrm{st}} &\geq \|i\cdot z\|_{\mathrm{st}} - \|z_i - i\cdot z\|_{\mathrm{st}} \geq i - r\,.\end{aligned}$$

Thus, after removing finitely many $z_i$ we will have $z_i \neq 0$ and $\lim_{i\to\infty} i/\|z_i\|_{\mathrm{st}} = 1$. Therefore, we obtain

$$\lim_{i\to\infty} \frac{z_i}{\|z_i\|_{\mathrm{st}}} = \lim_{i\to\infty} \frac{i\cdot z}{i} + \lim_{i\to\infty} \frac{z_i - i\cdot z}{i} = z + 0\,.$$

Moreover, for all $i \in \mathbb{N}$, we have

$$\begin{aligned}\langle \omega, z_i\rangle - \|z_i\|_{\mathrm{st}} &\geq \langle \omega, i\cdot z\rangle + \langle \omega, z_i - i\cdot z\rangle - \|i\cdot z\|_{\mathrm{st}} - \|z_i - i\cdot z\|_{\mathrm{st}}\\ &\geq i - \|\omega\|^*_{\mathrm{st}} \cdot \|z_i - i\cdot z\|_{\mathrm{st}} - i - r\\ &\geq -2\cdot r\,.\end{aligned}$$

Because $N$ and $\|\bullet\|_{\mathrm{st}}$ are uniformly close (Theorem 4.1), the claim follows – with the constant $a := 2\cdot r + D$, where $D$ is provided by Theorem 4.1; a possible concrete choice for $D$ is given by Equation (4.4). $\square$

Let $(z_i)_{i\in\mathbb{N}} \subset H_1(M;\mathbb{Z})_\mathbb{R} \setminus \{0\}$ and $a \in \mathbb{R}_{\geq 0}$ be as provided by Lemma 6.1. As next step, we choose corresponding 1-forms:

**Lemma 6.2.** *Choose a number $\ell \in \{1,\dots,b\}$ such that $\langle [\alpha^\ell], z\rangle \neq 0$. In the situation described above, we can find a sequence $(\lambda_i)_{i\in\mathbb{N}}$ in $\mathbb{R}$ with $\lim_{i\to\infty}\lambda_i = 0$ such that*

$$\eta_i := \eta + \lambda_i \cdot \alpha^\ell$$

*satisfies $\big\langle [\eta_i], z_i\big\rangle = 0$ for all $i \in \mathbb{N}$. In particular, each $\eta_i$ is an $\mathbb{R}$-linear combination of $\alpha^1,\dots,\alpha^b$.*

*Proof.* For $i \in \mathbb{N}$, we abbreviate $\overline{z}_i := z_i/\|z_i\|_{\mathrm{st}}$ and write

$$\overline{z}_i = \sum_{j=1}^b \mu_i^j \cdot \beta_j \quad \text{and} \quad z = \sum_{j=1}^b \mu^j \cdot \beta_j$$

in the chosen basis $(\beta_1,\dots,\beta_b)$ of $H_1(M;\mathbb{R})$. Then $(\mu_i^j)_{i\in\mathbb{N}}$ converges to $\mu^j$. As $([\alpha^1],\dots,[\alpha^b])$ is the dual basis we have $\mu^j = \langle [\alpha^j], z\rangle$, and thus by assumption $\mu^\ell \neq 0$. Passing to a subsequence, we may achieve $|\mu_i^\ell| \geq |\mu^\ell|/2 > 0$ for all $i \in \mathbb{N}$. For $i \in \mathbb{N}$, we set

$$\lambda_i := -\frac{1}{\mu_i^\ell} \cdot \langle [\eta], \overline{z}_i\rangle\,.$$

This sequence has the desired properties: Because the sequence $(|1/\mu_i^\ell|)_{i\in\mathbb{N}}$ is bounded by $\leq 2/|\mu^\ell|$ and $\lim_{i\to\infty}\langle[\eta],\overline{z}_i\rangle = \langle[\eta],z\rangle = 0$, we obtain $\lim_{i\to\infty}\lambda_i = 0$. Moreover, for all $i\in\mathbb{N}$, duality of the bases leads to

$$\begin{aligned}\big\langle[\eta_i],\overline{z}_i\big\rangle &= \big\langle[\eta],\overline{z}_i\big\rangle + \lambda_i\cdot\Big\langle[\alpha^\ell],\sum_{j=1}^b \mu_i^j\cdot\beta_j\Big\rangle\\ &= \big\langle[\eta],\overline{z}_i\big\rangle + \sum_{j=1}^b \lambda_i\cdot\delta_{j\ell}\cdot\mu_i^j\\ &= \big\langle[\eta],\overline{z}_i\big\rangle + \lambda_i\cdot\mu_i^\ell\\ &= 0\,. \end{aligned}$$

□

Our goal now is to construct new minimal geodesics. Additionally we also want to show that these geodesics are $\mathbb{R}$-homologically minimal geodesics. We will achieve this by constructing a geodesic $\mathbb{R}\to\widehat{M}^{\mathbb{R}}$ that globally minimizes the distance. Readers of this construction may replace $\widehat{M}^{\mathbb{R}}$ by the universal covering $\widetilde{M}$ in their first reading, as this case might be more transparent and already allows to understand why the constructed geodesics are minimal (without the stronger conclusion of $\mathbb{R}$-homological minimality).

From now on, let $(\eta_i)_{i\in\mathbb{N}}$ be as provided by Lemma 6.2. For each $i\in\mathbb{N}$, we choose a closed smooth geodesic $\sigma_i$ (parametrised by arclength) of length $\ell_i := N(z_i)$ representing $z_i$ and a lift $\widehat{\sigma}_i\colon\mathbb{R}\to\widehat{M}^{\mathbb{R}}$ to $\widehat{M}^{\mathbb{R}}$. By Theorem 4.1 and Lemma 6.1, we have $\ell_i = N(z_i)\geq\|z_i\|_{\mathrm{st}}\to\infty$. The function

$$u_i\colon\mathbb{R}\to\mathbb{R}.\quad t\mapsto\int_{\sigma_i|_{[0,t]}}\eta_i \tag{6.1}$$

is smooth and $\ell_i$-periodic; indeed, for all $t\in\mathbb{R}$, we have

$$u_i(t+\ell_i)-u(t)=\int_{\sigma_i|_{[t,t+\ell_i]}}\eta_i=\big\langle[\eta_i],z_i\big\rangle=0\,.$$

By reparametrizing $\sigma_i$ (and $\widehat{\sigma}_i$), we can achieve that $u_i$ attains its minimum at $t=0$. In particular, $u_i(0)=0$ and $u_i\geq 0$.

As $M$ is compact and as $\widehat{M}^{\mathbb{R}}\to M$ is a normal covering, we may assume (by translating the lifts via deck transformations and passing to a suitable subsequence) that the sequence $(\dot{\widehat{\sigma}}_i(0))_{i\in\mathbb{N}}$ converges to some unit vector $v_\infty\in T\widehat{M}^{\mathbb{R}}$. We then consider the limiting curve

$$\begin{aligned}\widehat{\sigma}_\infty\colon\mathbb{R}&\to\widehat{M}^{\mathbb{R}}\\ t&\mapsto\exp_{\widehat{M}^{\mathbb{R}}}(t\cdot v_\infty)\,.\end{aligned}$$

The curve $\widehat{\sigma}_\infty$ is a geodesic and the curves $\widehat{\sigma}_i$ converge on all compact intervals uniformly in the $C^\infty$-topology to $\widehat{\sigma}_\infty$ (Lemma 2.3, which is applicable in view of Remark 3.6). Thus, $\widehat{\sigma}_\infty$ minimizes distance between any of its points. Let $\sigma_\infty\colon\mathbb{R}\to M$ be the projection of $\widehat{\sigma}_\infty$ to $M$. By construction, $\sigma_\infty$ is a minimal geodesic on $M$ and the $\sigma_i$ converge on all compact intervals uniformly to $\sigma_\infty$. We say that $\sigma_\infty$ is a *minimal geodesic constructed from* $([x,y],z,\eta)$.

By construction, this is even an $\mathbb{R}$-homologically minimal geodesic.

6.3. **Asymptotes of geodesics from exposed edges of the stable norm ball.** In order to distinguish geodesics constructed from exposed edges as in Section 6.2, we analyze their asymptotes with respect to the underlying exposed edge of the stable norm unit ball. We recall that $A^{+\cup-}(\sigma_\infty) = A^+(\sigma_\infty) \cup A^-(\sigma_\infty)$ (Definition 5.1).

**Proposition 6.3.** *We assume the situation of Section 6.2. In particular, let $\sigma_\infty$ be a minimal geodesic constructed from $([x,y],z,\eta)$. Then we have*

$$A^{+\cup-}(\sigma_\infty) \subset [x,y]\,.$$

*Proof.* The cases of initial and terminal asymptotes are symmetric; we therefore only consider terminal asymptotes. Let $v \in A^+(\sigma_\infty)$. Again choose $\omega \in H^1_{\mathrm{dR}}(M)$ with $[x,y] = F_\omega$. Because $A^+(\sigma_\infty)$ lies in the stable norm unit ball (Lemma 5.4 (2)), it suffices to show that $\langle \omega, v\rangle = 1$.

On the one hand, by definition of $\|\bullet\|^*_{\mathrm{st}}$ we have $\langle\omega,v\rangle \leq \|\omega\|^*_{\mathrm{st}} \cdot \|v\|_{\mathrm{st}} \leq 1$. For the converse estimate, we argue as follows: Let $(t_i)_{i\in\mathbb{N}_0} \subset \mathbb{R}$ be a sequence of parameters with $t_i \to \infty$ realizing $v$ in Equation (5.1) of Definition 5.1; without loss of generality, we may assume that $t_0 = 0$ (Remark 5.2). Passing to a subsequence of $(\sigma_i)_{i\in\mathbb{N}}$ without changing the sequence $(t_i)_{i\in\mathbb{N}}$, we may assume that $t_i \in [0,\ell_i]$ for all $i \in \mathbb{N}$ and

$$\big\|h\big(\sigma_i|_{[0,t_i]}\big) - h\big(\sigma_\infty|_{[0,t_i]}\big)\big\|_{\mathrm{st}} \leq \frac{1}{i}\cdot t_i\,.$$

In particular, this implies that

$$\langle\omega,v\rangle = \lim_{i\to\infty}\left\langle \omega, \frac{h\big(\sigma_\infty|_{[0,t_i]}\big)}{t_i}\right\rangle = \lim_{i\to\infty}\left\langle \omega, \frac{h\big(\sigma_i|_{[0,t_i]}\big)}{t_i}\right\rangle$$

and we can thus exploit the properties of the geodesics $\sigma_i$. For all $i \in \mathbb{N}$, we obtain

$$\begin{aligned}\left\langle \omega, \frac{h\big(\sigma_i|_{[0,t_i]}\big)}{t_i}\right\rangle &= \frac{1}{t_i}\cdot\Big(\big\langle\omega, h\big(\sigma_i|_{[0,\ell_i]}\big)\big\rangle - \big\langle\omega, h\big(\sigma_i|_{[t_i,\ell_i]}\big)\big\rangle\Big)\\ &\geq \frac{1}{t_i}\cdot\Big(\langle\omega,z_i\rangle - \underbrace{\|\omega\|^*_{\mathrm{st}}}_{=1}\cdot\big\|h\big(\sigma_i|_{[t_i,\ell_i]}\big)\big\|_{\mathrm{st}}\Big)\,.\end{aligned}$$

Proposition 4.6 provides $-\big\|h\big(\sigma_i|_{[t_i,\ell_i]}\big)\big\|_{\mathrm{st}} \geq -\mathcal{L}(\sigma) - c$ for $c := (\mathcal{J}+1)\operatorname{diam} M$ and $\mathcal{J}$ given by Equation (4.2). Using this and Lemma 6.1, we continue our estimate with

$$\begin{aligned}\left\langle \omega, \frac{h\big(\sigma_i|_{[0,t_i]}\big)}{t_i}\right\rangle &\geq \frac{1}{t_i}\cdot\big(\langle\omega,z_i\rangle - (\ell_i - t_i) - c\big)\\ &\geq \frac{1}{t_i}\cdot\big((\ell_i - a) - (\ell_i - t_i) - c\big)\\ &= \frac{t_i - a - c}{t_i}\,.\end{aligned}$$

We conclude that

$$\langle\omega,v\rangle \geq \lim_{i\to\infty}\frac{t_i - a - c}{t_i} = 1\,,$$

as desired. $\square$

More precisely, we know in which parts of the exposed edge the terminal/initial asymptotes are located with respect to the point $z$:

**Proposition 6.4.** *We assume the situation of Section 6.2. In particular, let $\sigma_\infty$ be a minimal geodesic constructed from $([x,y],z,\eta)$. Then we have*

$$A^+(\sigma_\infty) \subset [z,y] \quad \textit{and} \quad A^-(\sigma_\infty) \subset [x,z]\,.$$

*Proof.* We prove the claim for $A^+(\sigma_\infty)$, the other case being symmetric. We know that $A^+(\sigma_\infty) \subset [x,y]$ by Proposition 6.3. Therefore, it suffices to show that $\langle [\eta], v\rangle \geq 0$ for all $v \in A^+(\sigma_\infty)$.

We consider the function

$$u_\infty \colon \mathbb{R} \to \mathbb{R}, \quad t \mapsto \int_{\sigma_\infty|_{[0,t]}} \eta\,.$$

Because $\eta$ is a linear combination of $\alpha^1, \dots, \alpha^b$, Remark 2.8 tells us that

$$u_\infty(t) = \langle [\eta], h\big(\sigma_\infty|_{[0,t]}\big)\rangle\,.$$

Below, we will show that $u_\infty \geq 0$. Assuming this estimate, we can argue as follows: Let $v \in A^+(\sigma_\infty)$ and let $(t_i)_{i\in\mathbb{N}} \subset \mathbb{R}_{>0}$ be a sequence for which we obtain $v$ as a limit as in Definition 5.1. Then

$$\big\langle [\eta], v\big\rangle = \lim_{i\to\infty} \frac{\big\langle [\eta], h\big(\sigma_\infty|_{[0,t_i]}\big)\big\rangle}{t_i} = \lim_{i\to\infty} \frac{u_\infty(t_i)}{t_i} \geq 0\,,$$

as desired.

It remains to show that $u_\infty \geq 0$: Let $t \in \mathbb{R}$ and let $I \subset \mathbb{R}$ be a compact interval with $0, t \in \mathring{I}$. By construction, $\sigma_i|_I \to \sigma_\infty|_I$ in $C^\infty(I, M)$ for $i \to \infty$. We assume that an $\ell \in \{1, \dots, b\}$ and a sequence $(\lambda_i)_{i\in\mathbb{N}}$ is chosen as in Lemma 6.2; this provides $\eta_i$ and then $u_i$ is defined in Equation (6.1). Furthermore, we have

$$\lim_{i\to\infty} \int_{\sigma_i|_{[0,t]}} \alpha^\ell = \int_{\sigma_\infty|_{[0,t]}} \alpha^\ell$$

and thus $\big(\int_{\sigma_i|_{[0,t]}} \alpha^\ell\big)_{i\in\mathbb{N}}$ is bounded. Therefore,

$$\begin{aligned} u_\infty(t) &= \int_{\sigma_\infty|_{[0,t]}} \eta = \lim_{i\to\infty} \int_{\sigma_i|_{[0,t]}} (\eta_i - \lambda_i \cdot \alpha^\ell) \\ &= \lim_{i\to\infty} \underbrace{\int_{\sigma_i|_{[0,t]}} \eta_i}_{=u_i(t)} - \lim_{i\to\infty} \Bigg( \underbrace{\lambda_i}_{\to 0} \cdot \underbrace{\int_{\sigma_i|_{[0,t]}} \alpha^\ell}_{\text{bounded}} \Bigg) \\ &= \lim_{i\to\infty} u_i(t) \geq 0\,. \end{aligned}$$

□

**Proposition 6.5** (homologically homoclinic minimal geodesics)**.** *In the situation of Section 6.2, let us assume that $\sigma_\infty$ has a unique terminal and a unique initial asymptote and that these coincide. Then this asymptote is the given edge element $z$.*

*Proof.* Because of $[x,z] \cap [z,y] = \{z\}$, this is immediate from Proposition 6.4. □

In general, we cannot expect to be in this situation. We always have the following, which will help us to distinguish minimal geodesics for different exposed edges:

**Proposition 6.6.** *In the situation of Section 6.2, the geodesic $\sigma_\infty$ constructed above satisfies*

$$A_{\mathrm{mix}}(\sigma_\infty) \cap (x,y) \neq \emptyset\,.$$

*Proof.* We know that $A^{+\cup-}(\sigma_\infty) \subset [x,y]$ (Proposition 6.3) and that $A^{+\cup-}(\sigma_\infty) \neq \emptyset$ (Lemma 5.4 (4)).

If $A^{+\cup-}(\sigma_\infty)$ contains more than one point, then we use the connectedness of $A_{\mathrm{mix}}(\sigma_\infty)$, which contains $A^{+\cup-}(\sigma_\infty)$ (Lemma 5.4 (8)). Then $A_{\mathrm{mix}}(\sigma_\infty) \cap (x,y)$ is non-empty because $[x,y]$ is an exposed edge.

Therefore, it suffices to consider the case that $A^{+\cup-}(\sigma_\infty)$ consists of a single point of the stable norm unit ball. In this case, we obtain $A^{+\cup-}(\sigma_\infty) = \{z\}$ from Proposition 6.5; moreover, $z$ lies in $(x,y)$. □

We obtain two geometrically distinct geodesics constructed from a single exposed edge:

**Corollary 6.7.** *Let $M$ be a closed connected Riemannian manifold and let $[x,y]$ be an exposed edge of the stable norm unit ball of $M$. Then there exist at least two geometrically distinct $\mathbb{R}$-homologically minimal geodesics $\gamma\colon \mathbb{R} \to M$ of $M$ with*

$$\begin{aligned}&A^+(\gamma) \text{ and } A^-(\gamma) \text{ are closed connected subsets of } [x,y]\\ &\text{and } A_{\mathrm{mix}}(\gamma) \cap (x,y) \neq \emptyset\,.\end{aligned} \tag{6.2}$$

*Proof.* Let $z \in (x,y)$. We choose $\eta \in \Omega^1(M)$ as in Section 6.2. Let $\sigma_\infty$ be a minimal geodesic constructed from $([x,y],z,\eta)$ (Section 6.2). Flipping the sign, let $x' := y$, $y' := x$, $z' := z$, and $\eta' := -\eta$. Then the construction from Section 6.2 also provides us with a minimal geodesic $\sigma'_\infty$ constructed from $([y,x],z,-\eta)$.

We obtain (Proposition 6.4 and Proposition 6.6)

$$\begin{aligned}A^+(\sigma_\infty) &\subset [z,y], & A^-(\sigma_\infty) &\subset [x,z], & A_{\mathrm{mix}}(\sigma_\infty) \cap (x,y) &\neq \emptyset,\\ A^+(\sigma'_\infty) &\subset [x,z], & A^-(\sigma'_\infty) &\subset [z,y], & A_{\mathrm{mix}}(\sigma'_\infty) \cap (x,y) &\neq \emptyset\,.\end{aligned}$$

As $[x,y] \subset \partial B$ is an exposed edge, we have $[x,y] \cap [-x,-y] = \emptyset$. In particular, Remark 5.3 yields that $\sigma_\infty$ and $\sigma'_\infty$ are geometrically distinct unless

$$A^+(\sigma_\infty) = \{z\} = A^-(\sigma_\infty)\,. \tag{6.3}$$

If this holds (i. e., if $\sigma_\infty$ is homologically homoclinic), then we may pick $z'' \in (z,y)$ and an appropriate 1-form $\eta''$ and construct a minimal geodesic $\sigma''_\infty$ constructed from $([x,y],z'',\eta''\rangle)$. This minimal geodesic satisfies (Proposition 6.4 and Proposition 6.6)

$$A^+(\sigma''_\infty) \subset [z'',y] \quad \text{and} \quad A_{\mathrm{mix}}(\sigma''_\infty) \cap (x,y) \neq \emptyset\,.$$

This implies $z \in A^+(\sigma_\infty) \setminus A^+(\sigma''_\infty)$, and because of $[x,y] \cap [-x,-y] = \emptyset$ we conclude that $\sigma_\infty$ and $\sigma''_\infty$ are geometrically distinct.

By construction $\sigma_\infty$, $\sigma'_\infty$, and $\sigma''_\infty$ are $\mathbb{R}$-homologically minimal. The sets $A^\pm(\sigma_\infty)$, $A^\pm(\sigma'_\infty)$, and $A^\pm(\sigma''_\infty)$ are closed and connected by Lemma 5.4 (5) and (6). □

A stronger version of Corollary 6.7 is provided by the following:

**Proposition 6.8.** *Let $M$ be a closed connected Riemannian manifold and let $[x,y]$ be an exposed edge of the stable norm unit ball of $M$; and let $x$ and $y$ be exposed points of the stable norm unit ball. Then at least one of the following statements holds:*

(A) *There are uncountably many geometrically distinct homologically homoclinic minimal geodesics of $M$ satisfying Equation (6.2).*
(B) *There are infinitely many geometrically distinct homologically non-homoclinic minimal geodesics of $M$ satisfying Equation (6.2).*

(C) *There are geometrically distinct homologically non-homoclinic minimal geodesics* $\gamma_1, \dots, \gamma_m \colon \mathbb{R} \to M$ *satisfying Equation* (6.2) *with at least one of the following additional properties:*
 (C.i) $m = 2$, *and* $\gamma_1$ *and* $\gamma_2$ *are homologically heteroclinic and homologically exposed.*
 (C.ii) $m = 3$, *and* $\gamma_1$ *is homologically heteroclinic and homologically exposed. The geodesics* $\gamma_2$ *and* $\gamma_3$ *are homologically semi-exposed, but not homologically exposed.*
 (C.iii) $m = 4$, *and* $\gamma_1, \dots, \gamma_4$ *are homologically semi-exposed, but not homologically exposed.*

*All minimal geodesics detected in the above cases are* $\mathbb{R}$*-homologically minimal. Furthermore each of these geodesics is homologically non-exposed or it is homologically heteroclinic.*

**Remark 6.9.** The following proof remains valid if we remove the assumption that the boundary points $x$ and $y$ are exposed. Note that the exposedness of $[x, y]$ already implies that $x$ and $y$ are extremal, which is a bit weaker than being exposed, see Subsection 3.4. If $x$ is no longer exposed, then $A^{\pm}(\gamma) = \{x\}$ no longer means that $\gamma$ is "homologically semi-exposed", see Definition 5.13, and similarly for $y$. Nevertheless Proposition 6.8 still holds if we replace the properties "homologically exposed", "homologically semi-exposed" in Proposition 6.8 (C) by "homologically extremal", "homologically semi-extremal" which are defined similarly by replacing "exposed points" by "extremal points" in Definition 5.13. We will not elaborate on this in detail, as the main focus of the article lies on the case that the stable norm unit ball is a polytope in which case all faces of this ball are exposed, see the comment at the end of Subsection 3.4.

This proposition and its proof show, in particular: If the stable norm unit ball contains an exposed edge $[x, y]$, then

- Condition (A) holds or
- the two geodesics provided by Corollary 6.7 may be chosen to be homologically non-homoclinic.

For both (non-exclusive) alternatives, there are examples, e. g., Example 6.10 for the first alternative, and for the second alternative Examples 9.4.

*Proof of Proposition 6.8.* For simplicity of notation we introduce a linear ordering $\leq$, resp. $<$, on $[x, y]$, characterized by "$z \leq w \Leftrightarrow w \in [z, y]$". All minimal geodesics in this proof will be $\mathbb{R}$-homologically minimal.

The proof of this proposition starts as the proof of Corollary 6.7. For every given $z \in (x, y)$ we obtain $\mathbb{R}$-homologically minimal geodesics $\gamma_1^z \coloneqq \sigma_\infty$ and $\gamma_2^z \coloneqq \sigma'_\infty$ with

$$
\begin{aligned}
&A^+(\gamma_1^z) \subset [z, y], \quad A^-(\gamma_1^z) \subset [x, z], \quad A_{\mathrm{mix}}(\gamma_1^z) \cap (x, y) \neq \emptyset, \\
&A^+(\gamma_2^z) \subset [x, z], \quad A^-(\gamma_2^z) \subset [z, y], \quad A_{\mathrm{mix}}(\gamma_2^z) \cap (x, y) \neq \emptyset\,.
\end{aligned}
$$

As all $A^{\pm}(\gamma_i^z)$ are closed and connected subsets of $[x, y]$, there are $v_i^{\pm}(z) \in [x, y]$ and $w_i^{\pm}(z) \in [x, y]$ with $A^{\pm}(\gamma_i^z) = [v_i^{\pm}(z), w_i^{\pm}(z)]$, and we get

$$
\begin{aligned}
x \leq v_1^-(z) \leq w_1^-(z) \leq z \leq v_1^+(z) \leq w_1^+(z) \leq y\,, \\
x \leq v_2^+(z) \leq w_2^+(z) \leq z \leq v_2^-(z) \leq w_2^-(z) \leq y\,.
\end{aligned}
$$

We study the geodesics $\gamma_1^z$ with $z \in (x, y)$ first, which are not necessarily geometrically distinct for different values of $z$. We claim that in this family $(\gamma_1^z)_{z\in(x,y)}$ of geodesics

(a) there is at least one homologically heteroclinic and homologically exposed one,
(b) or there are at least two geometrically distinct homologically semi-exposed ones that are not homologically exposed and that are homologically non-homoclinic,
(c) or this family provides an infinite number of geometrically distinct homologically non-homoclinic minimal geodesics,
(d) or there are uncountably many geometrically distinct homologically homoclinic minimal geodesics.

If we carry out the same discussion for $\gamma_2^z$, the proposition follows.

We consider several cases:

*Case 1: For all $z \in (x, y)$, we have $v_1^+(z) < y$.*

In this case, we choose a $z_1 \in (x, y)$ and then we inductively select

$$z_{i+1} \in \bigl(v_1^+(z_i), y\bigr)\,.$$

By construction, the geodesics $\gamma_1^{z_i}$ with $i \in \mathbb{N}$ are geometrically distinct minimal geodesics. We have thus obtained an infinite number of minimal geodesics. If infinitely many of them are homologically non-homoclinic, then we have Item (c) of the claim. If only finitely many of them are geometrically non-homoclinic with Equation (6.2), then the union of their $A_{\mathrm{mix}}$-sets is a compact subset of $[x, y)$ and thus its complement contains a set of the form $[w, y)$ for some $w < y$. Then for every $z \in [w, y)$, the geodesic $\gamma_1^z$ is a homologically homoclinic minimal geodesic, and they are pairwise geometrically distinct. We thus get Item (d) of the claim.

*Case 2: For all $z \in (x, y)$, we have $x < w_1^-(z)$.*

In this case, we argue analogously to Case 1.

*Case 3: There exists a $z \in (x, y)$ with $x = w_1^-(z)$ and $v_1^+(z) = y$.*

We choose such a $z$. Then $A^+(\gamma_1^z) = \{y\}$ and $A^-(\gamma_1^z) = \{x\}$. The geodesic $\gamma_1^z$ is homologically heteroclinic and homologically exposed, and Item (a) of the claim holds.

*Case 4: Remaining case.*

In this case, we have

$$\begin{aligned} &\exists_{z\in(x,y)} \quad \bigl(x < w_1^-(z) \text{ and } v_1^+(z) = y\bigr) \quad \text{and} \\ &\exists_{z'\in(x,y)} \quad \bigl(x = w_1^-(z') \text{ and } v_1^+(z') < y\bigr)\,. \end{aligned}$$

The minimal geodesics $\gamma_1^z$ and $\gamma_1^{z'}$ are homologically semi-exposed and geometrically distinct. They are not homologically exposed. Thus, we have Item (b) of the claim. □

**Example 6.10.** Let $M = T^2 \times S^2$, where $T^2 = \mathbb{R}^2/\mathbb{Z}^2$. We choose the 1-forms $\alpha^1$ and $\alpha^2$ in Setup 2.4 as pullbacks of $\mathrm{d}x^1$ and $\mathrm{d}x^2$ on $\mathbb{R}^2/\mathbb{Z}^2$. The projection $p_1 \colon M \to T^2$ gives an isomorphism $H_1(M;\mathbb{R}) \to H_1(T^2;\mathbb{R})$ and we identify $H_1(M;\mathbb{R}) \cong H_1(T^2;\mathbb{R}) \cong \mathbb{R}^2 \cong \widetilde{T}^2$. The Jacobi maps $J^M$ and $J^{T^2}$ of $M$ and $T^2$ satisfy $J^M = J^{T^2} \circ \widetilde{p}_1$, where $\widetilde{p}_1 \colon \mathbb{R}^2 \times S^2 \to \mathbb{R}^2$ is again projection to the first factor.

Let $g_s$ be a Riemannian metric with constant coefficients on $\mathbb{R}^2$, depending smoothly on $s \in S^2$; we also write $g_s$ for the associated metric on $T^2$. Denoting the standard metrik on $S^2$ by $g^S$ we define the Riemannian metric $g \coloneqq g_s + g^S$ on $M$. One may choose a smooth family $f_s \in \mathrm{Iso}(\mathbb{R}^2)$, such that $f_s^* g_s$ is the standard metric on $\mathbb{R}^2$. The closed unit norm ball of $g_s$ is thus $B_s \coloneqq f_s(B_1(0))$, where $B_1(0)$ is the standard closed unit ball in $\mathbb{R}^2$.

Then $B = \bigcup_{s\in S^2} B_s$ is the union of a smooth family of ellipses centered in 0; however, in general, $B$ is not convex. But for suitable choices of $g_s$, one can achieve that $B$ is convex and such that there is an embedded smooth path $\sigma\colon [0,1] \to S^2$ with $B = \bigcup_{t\in[0,1]} B_{\sigma(t)}$. These conditions imply

$$\|v\|_{\mathrm{st}} = \min_{t\in[0,1]} \sqrt{g_{\sigma(t)}(v,v)}$$

and that $B$ is the stable norm unit ball. By choosing the family $(g_s)_s$ in a suitable way, we can additionally achieve that $\partial B$ contains an exposed edge, i. e., a non-trivial line segment $e \subset \partial B$. Let $\mathring{e} \neq \emptyset$ be the open line segment associated to $e$.

Under these assumptions we have on $(M, g)$:

- all minimial geodesics $\gamma$ with $A_{\mathrm{mix}}(\gamma) \subset e$ are homologically homoclinic,
- all points $v$ in $\mathring{e}$ have a minimal geodesic $\gamma_v$ with $A^+(\gamma_v) = \{v\} = A^-(\gamma_v)$ (we do not claim uniqueness of $\gamma_v$).

Obviously all minimal geodesics are $\mathbb{R}$-homologically minimal. Thus we have verified Condition (A).

## 7. Lower bounds for centrally symmetric polytopes

We give a (crude) lower bound for the number of vertices and edges of centrally symmetric polytopes, which will be used to prove Theorem 1.3 in Section 8.

**Proposition 7.1.** *Let $n \in \mathbb{N}$ and let $P \subset \mathbb{R}^n$ be a centrally symmetric polytope with non-empty interior. Let $V$ and $E$ denote the number of vertices and edges of $P$, respectively. Then, $V \geq 2\cdot n$ and*

$$\frac{1}{2}\cdot V + E \geq \min(n^2 + 2\cdot n + 1, 2\cdot n^2 - n) = \begin{cases} 2\cdot n^2 - n & \text{if } n \leq 3 \\ n^2 + 2\cdot n + 1 & \text{if } n \geq 4. \end{cases}$$

*Proof.* We distinguish two cases:

(1) The polytope $P$ is simplicial, i.e., all faces are simplices. Because $P$ is centrally symmetric, the number of vertices and edges of the $n$-dimensional cross-polytope is a lower bound for $V$ and $E$, respectively [30, Section 7]. Hence,

$$V \geq 2\cdot n \quad \text{and} \quad E \geq 2^{1+1}\cdot \binom{n}{1+1} = 2\cdot n\cdot (n-1)\,.$$

Therefore, $1/2\cdot V + E \geq 2\cdot n^2 - n$.

(2) The polytope $P$ is *not* simplicial, i.e., $P$ has at least one facet $\sigma$ (i.e., at least one face of dimension $n-1$ in the boundary) that is not an $(n-1)$-simplex. In particular, $\sigma$ has at least $n+1$ vertices. Because $P$ is centrally symmetric, also $-\sigma$ is a facet of $P$; as $P$ has non-empty interior, these parallel faces $\sigma$ and $-\sigma$ have no common vertex. Thus, $P$ has at least $V \geq 2\cdot(n+1)$ vertices.

Because of $\mathring{P} \neq \emptyset$, each vertex of $P$ has edge-degree at least $n$. Therefore, $P$ has at least

$$\frac{1}{2} \cdot n \cdot V \geq \frac{1}{2} \cdot n \cdot 2 \cdot (n+1) = n \cdot (n+1)$$

edges. Hence, $1/2 \cdot V + E \geq (n+1) \cdot (n+1) = n^2 + 2 \cdot n + 1$.

Taking the minimum over both cases gives the claim. □

We can reformulate this estimate in terms of the constants from Definition 1.2:

**Corollary 7.2.** *For all $b \in \mathbb{N}$, we have*

$$\mathrm{VE}_{\min}(b) \geq \begin{cases} 2 \cdot b^2 - b & \text{if } b \leq 3 \\ b^2 + 2 \cdot b + 1 & \text{if } b \geq 4. \end{cases}$$

*Furthermore,* $\mathrm{VE}_{\min}(b) \geq b + \mathrm{E}_{\min}(b)$*; if* $\mathrm{VE}_{\min}(b) = b + \mathrm{E}_{\min}(b)$*, then* $\mathrm{E}_{\min}(b)$ *is attained by the cross-polytope, and we thus have* $\mathrm{E}_{\min}(b) = 2 \cdot b \cdot (b-1)$.

*Proof.* The first lower bound on $\mathrm{VE}_{\min}(b)$ is a direct consequence of Proposition 7.1. Let $P$ be a polytope that attains $\mathrm{VE}_{\min}(b)$. By elementary geometry, $P$ has at least $2b$ vertices, which implies that $\mathrm{VE}_{\min}(b) \geq b + \mathrm{E}_{\min}(b)$. If this inequality is an equality, then $P$ has precisely $2b$ vertices, and is then (up to an affine transformation) a cross-polytope. □

In view of the square and the octahedron, these bounds are sharp in dimensions 2 and 3. Improved bounds for $\mathrm{E}_{\min}$ in higher dimensions are discussed in Remark 1.4 (a).

## 8. Proofs of the results from the introduction

We combine the results from the previous sections to complete the proofs of the main results. We will always use the following notation:

**Setup 8.1.** Let $M$ be a closed connected (and non-empty) Riemannian manifold. Let $b := \dim_{\mathbb{R}} H_1(M;\mathbb{R})$ and let $B \subset H_1(M;\mathbb{R})$ be the stable norm unit ball.

**Remark 8.2.** We know that $B$ is closed, convex, centrally symmetric, and has non-empty interior. If $B$ has only finitely many exposed points, then $B$ has only finitely many extreme points (by Straszewicz's theorem [34]) and so $B$ is a compact centrally symmetric polytope in $H_1(M;\mathbb{R})$ with non-empty interior.

8.1. **Proof of Theorem 1.3.** Let $M$ be a closed connected Riemannian manifold. The stable norm ball $B \subset H_1(M;\mathbb{R})$ is centrally symmetric and convex.

(1) If $v$ is an exposed point of $B$, then Bangert shows that there exists an $\mathbb{R}$-homologically minimal geodesic $\gamma$ with $A^+(\gamma) = \{v\} = A^-(\gamma)$ [6, Theorem 4.4]. This proves the theorem in the case of $b = 1$ or in the case that $B$ has infinitely many exposed points.

(2) We may therefore assume that $b \geq 2$ and that $B$ has only finitely many exposed points. Then $B$ is a compact centrally symmetric polytope with non-empty interior (Remark 8.2); in particular, all vertices/edges of $B$ are exposed. Let $V(B)$ denote the number of vertices of $B$ and $E(B)$ the number of edges of $B$.

Applying Bangert's construction to each antipodal pair of vertices of $B$ provides $1/2 \cdot V(B)$ geometrically distinct minimal geodesics (which are homologically homoclinic and homologically exposed).

For each exposed edge $[x, y]$ of $B$, we obtain two geometrically distinct minimal geodesics $\gamma$ with $A_{\text{mix}}(\gamma) \cap (x, y) \neq \emptyset$ (Corollary 6.7). In particular, such minimal geodesics for different antipodal edge pairs are geometrically distinct. Thus we have at least $E(B)$ many geometrically distinct minimal geodesics of this type. Moreover, these minimal geodesics for exposed edges do not have the vertex asymptote behaviour of Bangert's vertex examples: they are homologically heteroclinic or they are not homologically exposed. Thus the geodesics constructed from the edges are all geometrically distinct from the ones obtained from the vertices.

The total number of $\mathbb{R}$-homologically minimal geodesics is therefore at least $1/2 \cdot V(B) + E(B)$.

(3) Theorem 1.3 now follows from Corollary 7.2.

Moreover, we record the following observation from the proof above:

**Remark 8.3.** If $M$ has only finitely many geometrically distinct minimal geodesics, then $B$ is a compact convex polytope in $H_1(M; \mathbb{R})$ with non-empty interior and $M$ has at least $V(B)/2 + E(B)$ geometrically distinct minimal geodesics.

8.2. **Proof of Theorem 1.6.** If $B$ is not a polytope, then $B$ has infinitely many different antipodal pairs of exposed points (Remark 8.2) and so the corresponding minimal geodesics by Bangert show that we are in case (a) of Theorem 1.6.

If $B$ is a polytope, then the arguments in Section 8.1 can be refined. For every edge we apply Proposition 6.8:

- If we are in case (A) of Proposition 6.8 for some edge, then we have uncountably many geometrically distinct homologically homoclinic minimal geodesics. In particular, we are again in case (a) of Theorem 1.6.
- If we are in case (B) of Proposition 6.8 for some edge, then we have infinitely many geometrically distinct homologically non-homoclinic minimal geodesics. These geodesics are geometrically distinct from at least $b$ homologically homoclinic minimal geodesics, detected with Bangert's method. This shows that we are in case (b) of Theorem 1.6.
- If we are in case (C) of Proposition 6.8 for every edge, then every pair of antipodal edges contributes at least two homologically non-homoclinic minimal geodesics. The statament for the geodesics via Bangert's method is as above. Again, we are in case (b) of Theorem 1.6.

As before, all minimal geodesics in this list are $\mathbb{R}$-homologically minimal.

The last claim of Theorem 1.6 follows from Remark 8.3 and the proof is thus complete.

8.3. **Proof of Theorem 1.7.** Let $M$ have exactly $b + \mathrm{E}_{\min}(b)$ minimal geodesics. In particular, in combination with Remark 8.3, we obtain that the stable norm unit ball $B$ is a compact convex centrally symmetric polytope in $H_1(M; \mathbb{R})$ with non-empty interior and $E(B) = \mathrm{E}_{\min}(b)$ as well as $V(B) = 2 \cdot b$. As vertices come in antipodal pairs, there is a $b$-tuple $\mathcal{V} = (v_1, \dots, v_b)$ of vectors in $H_1(M; \mathbb{R})$ such that $v_1, \dots, v_b, -v_1, \dots, -v_b$ are precisely the vertices of $B$. Because $B$ is the convex hull of its vertices and because $B$ has non-empty interior, $\mathcal{V}$ is a linearly independent family, and thus a basis of $H_1(M;\mathbb{R})$. It follows that $B$ has the combinatorial type of a cross-polytope.

In particular, $B$ is a simplicial polytope and as in the proof of Proposition 7.1 we have $\mathrm{E}_{\min}(b) = E(B)$. Therefore,

$$\mathrm{E}_{\min}(b) = E(B) = 2 \cdot b^2 - 2 \cdot b\,.$$

Further, every antipodal pair of edges contributes at least two $\mathbb{R}$-homologically minimal geodesics (Proposition 6.8) and every antipodal pair of vertices contributes one minimal geodesic. Thus, the equality discussion implies that each edge pair contributes precisely two, and each vertex pair precisely one. Hence, for the edges we will always be in case (C.i) of this proposition. It follows that we have $\mathrm{E}_{\min}(b)$ of $\mathbb{R}$-homologically minimal geodesics that are homologically heteroclinic and homologically exposed, and $b$ of $\mathbb{R}$-homologically minimal geodescis that are homologically homoclinic and homologically exposed, and that there are no other minimal geodesics (up to geometric equivalence).

8.4. **Proof of Theorem 1.8.** Let $M$ have exactly $V(B)/2 + E(B)$ geometrically distinct minimal geodesics. Then a similar equality discussion implies that precisely $E(B)$ minimal geodesics are determined by the edges and that precisely $V(B)/2$ minimal geodesics are determined by the vertices. The types of the minimal geodesics are then determined by Proposition 6.8, similarly to the proof of Theorem 1.7 in Subsection 8.3.

## 9. Examples

In the following section, we discuss existence of minimal geodesics on surfaces, as well as examples of Hedlund metrics on manifolds of dimension at least 3. We also discuss further examples and discuss more related literature.

9.1. **Surfaces.** We discuss minimal geodesics on closed connected surfaces $M$. On the one hand, the 2-dimensional case is special as topological conditions often imply intersections of geodesics and the results are in some sense different from higher dimensions, see Remark 9.3. On the other hand, this situation nicely illustrates possible behaviours of minimal geodesis and their asymptotes.

By passing to connected components and to the orientation covering, it is sufficient to consider connected and orientable surfaces $M$. Obviously, minimal geodesics exist if and only if the genus of $M$ is at least 1.

Let $M$ be such a surface and let $\gamma$ be a simple closed geodesic on $M$. Recall that *simple* means by definition that $\gamma$ defines an embedding of a circle $S^1$ into $M$. We also assume that $\gamma$ minimizes length within its free homotopy class. Such minimizing simple closed geodesics $\gamma$ always exist, in particular, in every primitive class in $H_1(M;\mathbb{Z})$ [25]. An element $\neq 0$ in a free abelian group is called *primitive* if it is not a non-trivial multiple of another element. We declare 0 to be *non-primitive.*

By a classical argument, carried out for the torus already by Hedlund [19, before Theorem II], the geodesic $\gamma$ is minimal; see Proposition C.2 in Appendix C for the general case.

**Example 9.1** (2-tori)**.** We consider the 2-dimensional torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, equipped with an arbitrary (smooth) Riemannian metric $g$. This example has many features essentially different from the higher dimensional case. Minimal geodesics on $T^2$ were already discussed in many aspects in articles by Hedlund [19] and Morse [28], where minimal geodesics are called "unending geodesics of class A". Some considerations even seem to go back to Bliss [9] who considered the special case of tori embedded in Euclidean $\mathbb{R}^3$ by rotating a circle in the $(x > 0, z \in \mathbb{R})$-half plane around the $z$-axis. We refer to Bangert's articles [5, 6] for a presentation in modern language.

We will again use the choices from Example 3.8; in particular, the Jacobi map corresponds to the identity under the canonical isomorphism $\mathbb{R}^2 \cong H_1(T^2;\mathbb{R})$. We identify $\pi_1(T^2) \cong H_1(T^2;\mathbb{Z}) \cong H_1(T^2;\mathbb{Z})_\mathbb{R} \cong \mathbb{Z}^2$, and elements thereof are written as pairs $(k,\ell)$ of integers. A pair is non-primitive if and only if it can be written as $(rk, r\ell)$, with $k,\ell,r \in \mathbb{Z}$, $r \geq 2$, and $(k,\ell) \neq 0$.

By an intersection argument, see Lemma C.1 in Appendix C for details, every closed curve representing a non-primitive non-zero element has to have a self-intersection, which implies that every closed geodesic of minimal length representing such a multiple $(rk, r\ell) \in H_1(T^2;\mathbb{Z})_\mathbb{R}$ is in fact an $r$-fold covering of a closed geodesic geodesic representing $(k,\ell)$. This implies that $N\big((rk,r\ell)\big) = r \cdot N\big((k,\ell)\big)$ and so $N\big((k,\ell)\big) = \|(k,\ell)\|_{\mathrm{st}}$.

If $(k',\ell') \notin \mathbb{Q} \cdot (k,\ell)$, then closed curves $\gamma$ and $\gamma'$ representing the homology classes $(k,\ell)$ and $(k',\ell')$, respectively, have to intersect. This implies

$$N\big((k+k', \ell+\ell')\big) < N\big((k,\ell)\big) + N\big((k',\ell')\big)\,.$$

We conclude that the stable norm unit sphere $\partial B$ does not contain any straight lines, i. e., the stable norm unit ball $B$ is strictly convex.

By definition, $\gamma\colon \mathbb{R} \to T^2$ is a minimal geodesic if a lift $\tilde{\gamma}\colon \mathbb{R} \to \mathbb{R}^2$ is a line, i. e., a globally minimizing geodesic, parametrized by arclength. It follows from Theorem 5.6, or alternatively from Hedlund's work, that for every minimal geodesic $\gamma$ there is a unique $v_\gamma \in H_1(T^2;\mathbb{R}) \cong \mathbb{R}^2$ with $\|v_\gamma\|_{\mathrm{st}} = 1$, such that

$$A^+(\gamma) = A^-(\gamma) = A_{\mathrm{mix}}(\gamma) = \big\{\lim_{t-s\to\infty} R\big(\gamma|_{[s,t]}\big)\big\} = \{v_\gamma\}\,.$$

Furthermore, $\tilde{\gamma}$ remains within bounded distance from the line $t \mapsto tv_\gamma$. Conversely, for every $v \in H_1(T^2;\mathbb{R})$ with $\|v\|_{\mathrm{st}} = 1$, there is a minimal geodesic $\gamma$ with $v_\gamma = v$ [6, Theorem 4.4]; however, it may happen that $v_\gamma = v_{\gamma'}$ for geometrically distinct minimal geodesics $\gamma$ and $\gamma'$.

We briefly indicate how *rotation* vectors are related to actual rotations.

There is a closed curve $\tau$ that has minimal length among all closed curves representing non-zero elements in $H_1(T^2;\mathbb{Z})$. One can show that this curve is a simple closed geodesic, and by the arguments above, it is also a minimal geodesic. Further, $[\tau] \in H_1(T^2;\mathbb{Z})$ is primitive; after pulling back by a diffeomorphism we may assume $[\tau] = (0,1)$. If $\gamma$ is a minimal geodesic with $v_\gamma = (v_1,v_2) \notin \mathbb{R}\cdot(0,1)$, then $v_\gamma$ and $\tau$ have infinitely many intersections, they all have the same orientation, and they are located at some $v_\gamma(t_i)$ for $i \in \mathbb{Z}$ with $t_i < t_{i+1}$. The map $v_\gamma(t_i) \mapsto v_\gamma(t_{i+1})$ defines a bijection $\varphi\colon A \to A$ for some subset $A$ of the circle $S \coloneqq \mathrm{image}(\tau)$, and $\varphi$ extends to a circle homeomorphism $\varphi\colon S \to S^1$ that lifts canonically to a periodic orientation-preserving homeomorphism $\Phi\colon \mathbb{R} \to \mathbb{R}$ [5]. Its rotation number

$$\rho(\Phi) \coloneqq \lim_{j\to\infty} \frac{1}{j} \cdot \big(\Phi^j(x) - x\big)$$

exists and does not depend on $x \in \mathbb{R}$ [29, Sec. 2.2]. One can show through appropriate choices that $\rho(\Phi) = v_2/v_1$. Thus the rotation vector $v_\gamma$ determines the rotation number $\rho(\Phi)$ and conversely the rotation vector $v_\gamma$, which is on the stable norm unit circle, is determined up to sign by $\rho(\Phi)$.

Recently, also the stable norm on slit 2-tori has been studied [27].

**Example 9.2.** Let $M$ be an orientable closed connected surface of genus $k \geq 2$ with an arbitrary Riemannian metric. Let $\gamma$ be a closed geodesic representing a non-trivial $\pi_1$-conjugacy class, see Section 2.2, contained in the commutator subgroup. Such geodesics exist, e. g., if $\langle a_1, \dots, a_k, b_1, \dots, b_k \mid [a_1, b_1] \cdot \dots \cdot [a_k, b_k] \rangle$ is a presentation of $\pi_1(M)$ arising from realizing $M$ in the standard way as a quotient of a $4k$-gon, then a curve minimizing $[a_1, b_1] \cdot \dots \cdot [a_\ell, b_\ell]$ with $\ell \in \{1, \dots, k-1\}$ in the associated free homotopy class has this property. It is not so straigthforward to decide whether $\gamma$ is a simple closed curve, but one can easily see that there are $\pi_1$-conjugacy classes and Riemannian metrics $g$ such that $\gamma$ is a simple closed curve. Let us mention here – we will not use this fact in this paper – that the simpleness of $\gamma$ is independent of the choice of Riemannian metric $g$, thus only depends on a suitably chosen $\pi_1$-conjugacy class [15, Theorem 2.1]. From the arguments at the beginning of this subsection, it follows that simple closed geodesics $\gamma$ are minimal geodesics. As $\gamma$ represents $0 \in H_1(M; \mathbb{Z})$, the set

$$\left\{ h(\gamma|_{[s,t]}) \;\middle|\; s, t \in \mathbb{R},\ s < t \right\}$$

is bounded in $H_1(M; \mathbb{R})$ and thus $A^{+\cup-}(\gamma) = A_{\mathrm{mix}}(\gamma) = \{0\}$. In particular, both sides of Equation (5.4) vanish. The associated periodically parametrized geodesic $\mathbb{R} \to M$ is a minimal geodesic, as explained above. Such geodesics are obtained in the work of Morse [28], Klingenberg [21], and Gromov [17, Section 7.5]. However, the minimal geodesics detected in Bangert's existence theorem [6] and in the present article satisfy $A^{+\cup-}(\gamma) \subset \{x \in H_1(M\ \mathbb{R}) \mid \|x\|_{\mathrm{st}} = 1\}$, see Proposition 5.5. Thus, the above geodesics will not arise this way.

**Remark 9.3.** The situation in dimension at least 3 is very different from the results sketched above. Topological conditions do no longer force intersections of curves, and this allows the construction of Hedlund examples, see Subsection 9.2. One of the consequences is that in dimension 2 the stable norm unit ball is never a polytope, see the discussion in an article by Massart [24, after Proposition 6] [23]. In dimension at least 3, however, Hedlund examples always exist and for such examples the stable norm unit ball is a polytope.

9.2. **Hedlund examples.** We now turn to the situation in higher dimensions. We briefly recall the construction of Hedlund examples on closed connected manifolds $M$ of dimension $n \geq 3$. The classical constructions of such metrics only treated the case of the $n$-dimensional torus with $n \geq 3$, the first – and name-giving – reference being Hedlund's work [19, Section 9]. The construction on the torus was also discussed in modified versions by Bangert [6, Section 5] and the first author [2]. To our knowledge, Hedlund metrics on arbitrary closed connected manifolds of dimension $\geq 3$ were constructed and discussed first in [1, IV.1.a and IV.1.b] under the name "Autobahnmetriken", also briefly sketched in [2, Sec. 6] as "Express-way metrics". Later they were discussed further [4, 20].

In order to construct a Hedlund metric on $M$, we take closed curves $\sigma_1, \dots, \sigma_\ell$ based in $x_0 \in M$, whose based homotopy classes $[\sigma_i]$ generate $\pi_1(M, x_0)$. One can easily achieve that these curves are closed regular curves (i. e., immersions of $S^1$), and because $\dim M = n \geq 3$ we may assume that the curves $\sigma_i$ are embeddings. Now we perturb the closed curves $\sigma_i \colon S^1 \to M$ in the class of free (i. e., unbased) closed curves to curves, denoted as $\tau_i \colon S^1 \to M$, which are embeddings of circles with pairwise disjoint images. One can easily construct a Riemannian metric $g_1$ on $M$ such that each $\tau_i$ is a closed geodesic of length 1 and such that the distance from $x_0$ to image$(\tau_i)$ is "well-controlled and sufficiently small".

Now let $L_1, \ldots, L_\ell \in \mathbb{R}_{>0}$ be given. We choose a smooth function $f\colon M \to \mathbb{R}$ such that

- $f\big|_{\mathrm{image}(\tau_i)} = L_i$, and $f > L_i$ in a tubular neighborhood of image$(\tau_i)$, away from image$(\tau_i)$,
- $f$ grows "rapidly" on a tubular neighborhood of the image$(\tau_i)$'s, and
- $f$ is "very large" suffiently far away from $\bigcup_{i=1}^{\ell}$ image$(\tau_i)$.
- It is also helpful to keep the value of $f$ controlled (and of medium size) close to chosen paths $\alpha_i$ from $x_0$ to the set image$(\tau_i)$.

One then defines the *Hedlund metric* $g_{\mathrm{Hed}} := f^2 g_1$. We refer to the literature [1, IV.1.a and IV.1.b] for a precise desciption of this metric for $\varepsilon := L_1 = L_2 = \ldots = L_\ell$, which extends to the generality sketched above.

The curves $\tau_i$ are then geodesics for $g_{\mathrm{Hed}}$. One can show that for suitable choices of $L_i$ the curves $\tau_i$ are minimal geodesics. Furthermore, every minimal geodesic $\gamma$ stays within a small tubular neighborhood of $\bigcup_{i=1}^{\ell}$ image$(\tau_i)$ "most of the time". The precise meaning of "most of the time" is a bit subtle. In the case that $\pi_1(M)$ is virtually nilpotent with the bounded minimal generation property [2, Def. 7.1][1, IV. Def 1.19] – which includes the case that $\pi_1(M)$ is abelian – "most of the time" means up to a bounded number of short intervals during which the geodesic $\gamma$ jumps from a neighborhood of some image$(\tau_i)$ to the neighborhood of some image$(\tau_j)$ with $j \neq i$ [1, IV.1. items c, e, f, g]. The geodesic $\gamma$ then gives rise to a "symbol sequence", which is a function

$$\mathbb{Z} \to \{\tau_1, \ldots, \tau_\ell, \overline{\tau}_1, \ldots, \overline{\tau}_\ell\}$$

unique up to shift in $\mathbb{Z}$; it describes the curves $\tau_i$ followed by $\gamma$ one after the other. See items (1)–(6) in the following example for a list of such sequences for the case $\pi_1(M) \cong \mathbb{Z}^2$.

Note that $h(\tau_1), \ldots, h(\tau_\ell)$ span $H_1(M;\mathbb{R})$ as a vector space. One can now show that for all $v \in H_1(M;\mathbb{R})$ we have

$$\|v\|_{\mathrm{st}} = \min\left\{ \sum_{i=1}^{\ell} |a_i| \cdot L_i \;\middle|\; a_1, \ldots, a_\ell \in \mathbb{R},\ v = \sum_{i=1}^{\ell} a_i \cdot h(\tau_i) \right\}.$$

If the geodesic is even $\mathbb{Z}$-homologically minimal, then the property to have only a "bounded number of jumps" holds for arbitrary fundamental groups. We consider this in more detail in some special cases.

**Example 9.4** (Hedlund examples with $b = 2$ and $B$ a parallelogram)**.** Let $M$ be a closed connected manifold of dimension at least 3 with $b = \dim H_1(M;\mathbb{R}) = 2$. Let $\tau_1$ and $\tau_2$ be two simple closed curves, representing a $\mathbb{Z}$-basis of $H_1(M;\mathbb{Z})_{\mathbb{R}} \cong \mathbb{Z}^2$. We choose a Hedlund metric $g_{\mathrm{Hed}}$ as in [1] or as above with $L_1 = L_2 = 1$. Then $H_1(M;\mathbb{R})$ with the stable norm is isometric to $\mathbb{R}^2$ with the $L^1$-norm with respect to the basis consisting of $x_{\mathbb{R}} := [\tau_1]$ and $y_{\mathbb{R}} := [\tau_2]$. We write $x_{\mathbb{Z}} := [\tau_1], y_{\mathbb{Z}} := [\tau_2]$ for the corresponding integral classes classes in $H_1(M;\mathbb{Z})$. With our conventions we have $-x_{\mathbb{Z}} = [\overline{\tau}_1]$ and $-y_{\mathbb{Z}} = [\overline{\tau}_2]$. As discussed above, every minimal geodesic $\gamma$ on $M$ with respect to the Hedlund metric $g_{\mathrm{Hed}}$ is described by a suitable word (of two-sided infinite length) over $\{x_{\mathbb{Z}}, y_{\mathbb{Z}}, x_{\mathbb{Z}}^{-1}, y_{\mathbb{Z}}^{-1}\}$. One can show that the following sequences give a full list of all symbol sequences of $\mathbb{R}$-homologically minimal geodesics (modulo geometric equivalence). In this list, a negative number $k$ has to be read as taking $|k|$ times the inverse:

(1) for $k \in \mathbb{Z}$: $\cdots x_{\mathbb{Z}} \cdots x_{\mathbb{Z}} \underbrace{y_{\mathbb{Z}} \cdots y_{\mathbb{Z}}}_{k \text{ times}} x_{\mathbb{Z}} \cdots x_{\mathbb{Z}} \cdots$
(2) for $k \in \mathbb{Z}$: $\cdots y_{\mathbb{Z}} \cdots y_{\mathbb{Z}} \underbrace{x_{\mathbb{Z}} \cdots x_{\mathbb{Z}}}_{k \text{ times}} y_{\mathbb{Z}} \cdots y_{\mathbb{Z}} \cdots$
(3) $\cdots x_{\mathbb{Z}} \cdots x_{\mathbb{Z}} y_{\mathbb{Z}} \cdots y_{\mathbb{Z}} \cdots$
(4) $\cdots y_{\mathbb{Z}} \cdots y_{\mathbb{Z}} x_{\mathbb{Z}} \cdots x_{\mathbb{Z}} \cdots$
(5) $\cdots x_{\mathbb{Z}} \cdots x_{\mathbb{Z}} \bar{y}_{\mathbb{Z}} \cdots \bar{y}_{\mathbb{Z}} \cdots$
(6) $\cdots \bar{y}_{\mathbb{Z}} \cdots \bar{y}_{\mathbb{Z}} x_{\mathbb{Z}} \cdots x_{\mathbb{Z}} \cdots$

Note that for each such symbol class there is an $\mathbb{R}$-homologically minimal geodesic. Furthermore, one can show – but this is supposedly not yet worked out in the literature – that for suitably chosen Hedlund metrics this minimal geodesic is uniquely determined (up to geometric equivalence) by its symbol sequence. If $\pi_1(M) \cong \mathbb{Z}^2$, then every minimal geodesic is $\mathbb{R}$-homologically minimal and thus this is the list of all minimal geodesics (up to geometric equivalence).

Minimal geodesics of type (1) and (2) are homologically homoclinic and homologically exposed. The case $k = 0$ for (1) resp. (2) describes $\tau_1$ and $\tau_2$. Bangert's existence results detect one minimal geodesic of type (1) and one of type (2).

The four minimal geodesics (3)–(6) are homologically heteroclinic and homologically exposed. Note that the stable norm is the $\ell^1$-norm with respect to this generating set. Thus, the stable norm unit ball is a polytope, more precisely a parallelogram. In the case that $b = 2$ and $B$ is a polytope, our existence theorem, Theorem 1.6 (b) predicts the existence of at least four homologically non-homoclinic and $\mathbb{R}$-homologically minimal geodesics. In the example above, these are the minimal geodesics (3)–(6). Thus, these examples show that our existence result provides an optimal lower bound – the number 4 – for homologically non-homoclinic and $\mathbb{R}$-homologically minimal geodesics.

The next example is an extension to this, in the sense that the case $k = 2$ in the following example reduces to the previous one.

**Example 9.5** (Hedlund examples with $b = 2$ and $B$ a $2k$-gon)**.** Let again $b = 2$ and $\dim M \geq 3$. Let $\tau_1, \dots, \tau_k \in H_1(M;\mathbb{Z})$ be given with images $\tau_1^{\mathbb{R}}, \dots, \tau_k^{\mathbb{R}}$ in $H_1(M;\mathbb{R})$. We further assume that the convex hull of $\{\tau_1^{\mathbb{R}}, \dots, \tau_k^{\mathbb{R}}, -\tau_1^{\mathbb{R}}, \dots, -\tau_k^{\mathbb{R}}\}$ is a symmetric set $B$ whose boundary is a convex $2k$-gon, and thus $E(B) = V(B) = 2k$. Then the Hedlund metric based on these curves $\tau_i$ with $L_i = 1$ has $B$ as stable norm unit ball. In this case, the last phrase in Theorem 1.6 predicts the existence of

- $2k$ homologically non-homoclinic and $\mathbb{R}$-homologically minimal geodesics, and
- $k$ homologically homoclinic and $\mathbb{R}$-homologically minimal geodesics.

This estimate is sharp: for the Hedlund metric in this example there are precisely $3k$ geodesics that are $\mathbb{R}$-homologically minimal; and among them there are $k$ homologically homoclinic ones, and $2k$ homologically heteroclinic (and thus homologically non-homoclinic) ones.

**Example 9.6** (homologically homoclinic or heteroclinic, but not homologically semi-exposed)**.** We describe the construction of a Hedlund metric with a finite number of $\mathbb{R}$-homologically minimal geodesics that are not homologically semi-exposed, but that are either homologically heteroclinic or homologically homoclinic. For simplicity of presentation, we specialize to the case of the 3-torus $M = T^3 = \mathbb{R}^3/\mathbb{Z}^3$, but similar constructions are also possible for every closed connected manifold $M$ with $n = \dim M \geq 3$ and $b = \dim_{\mathbb{R}} H_1(M;\mathbb{R}) \geq 2$.

Again as in Example 3.8 we assume that the Jacobi map $J$ corresponds to the identity from $\widetilde{T^3} = \widehat{T^3}^{\mathbb{Z}} = \widehat{T^3}^{\mathbb{R}} \cong \mathbb{R}^3$ to $H_1(T^3;\mathbb{R}) \cong \mathbb{R}^3$. In particular, the notion of "minimal geodesics" and of "$\mathbb{R}$- or $\mathbb{Z}$-homologically minimal geodesics" coincide. We make the choices such that $H_1(T^3;\mathbb{Z}) \cong H_1(T^3;\mathbb{Z})_{\mathbb{R}}$ is the lattice generated by the canonical basis $e_1, e_2, e_3$ of $\mathbb{R}^3$. For $i \in \{1,2,3\}$, we define $j(1) = 2$, $j(2) = 3$, $j(3) = 1$, and let

$$t \mapsto \tau_i(t) := \Big[ t \cdot e_i + \frac{1}{2} \cdot e_{j(i)} \Big] \in T^3 \quad \text{and} \quad L_i := 1\,.$$

Now we consider vectors $v_4 = \sum_{k=1}^3 v_4^k e_k$ and $v_5 = \sum_{k=1}^3 v_5^k e_k$ with $v_4^1, v_5^1 \in \mathbb{Z}$, $v_4^2, v_5^2 \in \mathbb{Z} \setminus \{0\}$, and $v_4^3, v_5^3 \in \mathbb{Z}_{>0}$; moreover, we set $L_i := \big|v_i^1\big| + \big|v_i^2\big| + \big|v_i^3\big|$. We assume that $v_i^1$, $v_i^2$, and $v_i^3$ have no common divisor $> 1$, i. e., $v_4$ and $v_5$ are primitive in $\mathbb{Z}^3$. Furthermore assume $v_4 \neq v_5$. We define for $i = 4, 5$

$$t \mapsto \tau_i(t) := \Big[ \frac{t}{L_i} \cdot v_i + w_i \Big] \in T^3\,,$$

where $w_4$ and $w_5$ are chosen such that $\tau_i$ and $\tau_j$ have disjoint images for $i, j \in \{1, \ldots, 5\}$ with $i < j$. All curves $\tau_1, \ldots, \tau_5$ are closed curves of periodicity $L_i$. We define $v_i := e_i$ for $i \in \{1, \ldots, 3\}$.

For these $\tau_i$ and $L_i$ and some basepoint $x_0 \in T^3 \setminus \bigcup_{i=1}^5 \operatorname{image}(\tau_i)$, we construct a Hedlund metric $g_{\mathrm{Hed}}$ as described above; let $d_{\mathrm{Hed}}$ be the associated distance function and let $D_{i,j} := \min\big\{ d_{\mathrm{Hed}}\big(\tau_i(t), \tau_j(s)\big) \mid t, s \in \mathbb{R} \big\} > 0$ for $i \neq j$. Let again $\alpha_i$ be a suitable path from $x_0$ to $\tau_i$ as in the construction of the function $f$ at the beginning of this subsection. We assume that $f$ is much smaller along $\alpha_4$ and $\alpha_5$ than along $\alpha_1, \alpha_2, \alpha_3$; and we also assume that outside tubular neighborboods of the $\tau_i$ and $\alpha_i$ the function $f$ is much larger. In particular, we may assume $3D_{4,5} < D_{i,j}$ for all $i \neq j$ with $\{i, j\} \neq \{4, 5\}$.

Then the stable norm on $H_1(T^3;\mathbb{R}) \cong \mathbb{R}^3$ is the $\ell^1$-norm with respect to $(e_1, e_e, e_3)$, and $\|v_i\|_{\mathrm{st}} = L_i$ for $i \in \{1, \ldots, 5\}$.

The closed curves $\tau_1, \ldots, \tau_5$ are minimal geodesics for $g_{\mathrm{Hed}}$. We have $A^+(\tau_i) = A^-(\tau_i) = \{(1/L_i)v_i\}$. Thus $\tau_1, \ldots, \tau_5$ are homologically homoclinic, and they are the only minimal geodesics of this type.

For $i \in \{1, \ldots, 3\}$, the unique initial and terminal asymptote $v_i = e_i$ of $\tau_i$ is an exposed point, thus $\tau_1, \tau_2, \tau_3$ are homologically exposed; furthermore, these minimal geodeiscs can be detected by Bangert's method. However, the unique asymptote of $\tau_4$ and $\tau_5$ is not exposed, and these $\tau_4$, $\tau_5$ are not homologically (semi-)exposed. They are minimal geodesics and they are homologically heteroclinic, but there are no homologically diverging minimal geodesics. For example, there is a minimal geodesic $\gamma$ with $A^-(\gamma) = \{(1/L_4)v_4\}$ and $A^+(\gamma) = \{(1/L_5)v_5\}$. Geodesics of such a type cannot be detected by Bangert's method. If $(1/L_4)v_4$ or $(1/L_5)v_5$ is not on an edge of the stable norm unit ball $B$, then this geodesic is also not detected by our method. The number of $\mathbb{R}$-homologically minimal geodesics is finite as previously shown by the first author [1, Korollar 1.16 in IV.1.e].

Our method provides the existence of homologically heteroclinic minimal geodesics with $A^-(\gamma) = \{\pm e_i\}$ and $A^+(\gamma) = \{\pm e_j\}$ with $i, j \in \{1, \dots, 3\}$ and $i \neq j$, where the two choices of sign are independent. This provides 12 geometrically distinct minimal geodesics.

Now let us consider the case that $(1/L_4)v_4$ lies on an edge of $B$, say in the relative interior of the edge $[e_2, e_3]$, e. g., $v_4 = e_2 + e_3$, $L_4 = 2$. We also assume $(1/L_5)v_5 \notin [e_2, e_3]$ Then it depends on some finer information about $f$ and some choices in our proof whether our method, applied to the edge $[e_2, e_3]$ will provide one of the following non-exclusive items (the number corresponds to the cases in Proposition 6.8)

(C.i) minimal geodesics $\gamma_1$ and $\gamma_2$ with

$$A^-(\gamma_1) = A^+(\gamma_2) = \{e_2\}, \quad A^+(\gamma_1) = A^-(\gamma_2) = \{e_3\}\,.$$

(C.ii) minimal geodesics $\gamma_1, \dots, \gamma_4$ with

$$A^-(\gamma_1) = A^+(\gamma_2) = \{e_2\}, \quad A^-(\gamma_3) = A^+(\gamma_1) = \{e_3\},$$
$$A^-(\gamma_4) = A^+(\gamma_4) = A^-(\gamma_2) = A^+(\gamma_3) = \left\{\frac{1}{L_4} v_4\right\}.$$

(Proposition 6.8 only provides $\gamma_1$ and two out of the three remaining ones.)

(C.iii) minimal geodesics $\gamma_1, \dots, \gamma_5$ with

$$A^-(\gamma_1) = A^+(\gamma_4) = \{e_2\}, \quad A^-(\gamma_3) = A^+(\gamma_2) = \{e_3\},$$
$$A^+(\gamma_5) = A^-(\gamma_5) = A^-(\gamma_2) = A^-(\gamma_4) = A^+(\gamma_1) = A^+(\gamma_3) = \left\{\frac{1}{L_4} v_4\right\}.$$

(Proposition 6.8 only provides the four semi-exposed geodesics $\gamma_1, \dots, \gamma_4$.)

9.3. **Some (expected) examples with minimal geodesics of other types.** Examples of Riemannian manifolds with homologically diverging minimal geodesics with abelian or nilpotent fundamental group are difficult to construct, see Remark 9.7. However, if the fundamental group is Gromov-hyperbolic, their existence follows from known facts, see Example 9.8.

Let us first formulate a candidate for a counterexample in Remark 9.7, which would imply that the bound in Proposition 6.8 (C.iii) cannot be improved to $m = 5$. Then we discuss the Gromov-hyperbolic case.

**Remark 9.7.** We expect that examples with the following properties exist, although to our knowledge the construction of such examples has not yet been worked out. These examples consist of a closed connected Riemannian manifold $(M, g)$, an exposed edge $[x, y]$ of the stable norm unit ball, and points $z_1 \in (x, y)$, $z_2 \in (z_1, y)$ such that there are *precisely* four geometrically distinct minimal geodesics $\gamma_1, \dots, \gamma_4 \colon \mathbb{R} \to M$ with $A_{\mathrm{mix}}(\gamma_i) \cap (x, y) \neq \emptyset$ such that

$$A^-(\gamma_1) = \{x\}, \quad A^+(\gamma_1) = [z_2, y], \quad A^-(\gamma_2) = [x, z_1], \quad A^+(\gamma_2) = \{y\}$$
$$A^+(\gamma_3) = \{x\}, \quad A^-(\gamma_3) = [z_2, y], \quad A^+(\gamma_4) = [x, z_1], \quad A^-(\gamma_4) = \{y\}\,.$$

Hence, these four geodesics would be homologically semi-exposed amd homologically semi-converging.

**Example 9.8.** Let $N$ be an oriented closed connected surface of genus $k \geq 2$ and let $M \coloneqq N \times S^2$. We choose a Hedlund metric on $M$, with respect to standard generators $a_1, \dots, a_k, b_1, \dots, b_k$ coming from writing $M$ as a quotient of a $4k$-gon in the standard way. This gives rise to disjoint simple closed curves $\tau_1, \dots, \tau_{2k}$, for

$L_i := 1$; we construct the associated Hedlund metric. Then $(H_1(M;\mathbb{R}), \|\bullet\|_{\mathrm{st}})$ is isometric to $\mathbb{R}^{2k}$ with the $\ell^1$-norm, and all $\tau_i$ are minimal geodesics. In every exposed edge of the stable norm unit ball, we find minimal geodesics $\gamma_i$, $i \in \{1, \ldots, 4\}$ with all the properties in Remark 9.7, but without the word "precisely". In the present example, infinitely many geometrically distinct minimal geodesics $\gamma$ with $A_{\mathrm{mix}}(\gamma) \subset [x, y]$ and $A_{\mathrm{mix}}(\gamma) \cap (x, y) \neq \emptyset$ do exist. Furthermore, we also can show the existence of homologically diverging, homologically semi-converging, homologicyally semi-exposed geodesics, and many more types.

9.4. **Comparison to results by Bolotin and Rabinowitz.** For the case that $M = T^n$ is an $n$-dimensional torus, Bolotin and Rabinowitz [10] shows the existence of homoclinic and heteroclinic minimal geodesics in the sense of Poincaré, see Subsection 5.5. These authors fix a non-trivial class $v \in H_1(T^n;\mathbb{Z}) \cong \pi_1(T_n) \cong \mathbb{Z}^n$, and two conditions (S1) and (S2) are assumed – see below for details.

Let $\mathcal{C}_v$ be the set of all $C^\infty_{\mathrm{pw}}$-loops representing $v$. For simplicity we identify loops in $\mathcal{C}_v$ that are reparametrizations of each other. According to Lemma A.2 the length functional $\mathcal{L}\colon \mathcal{C}_v \to [0, \infty)$ attains its infimum, and each minimizer $\tau \in \mathcal{C}_v$ is a closed geodesic (if parametrized proportionally to arclength). The results in the article [10] assume that at least one minimizer – and thus all minimizers – are minimal geodesics. Let us rescale the metric $g$ on $T^n$ such that $\|v\|_{\mathrm{st}} = 1$, i. e., for every minimizer $\tau$ we have $\mathcal{L}(\tau) = N([\tau]) = \|v\|_{\mathrm{st}} = 1$.

Assumption (S2) [10] claims that the number $p$ of minimizers in $\mathcal{C}_v$ is finite. We omit the precise definition of (S1), but let us mention that it follows from some mild conditions, e.g. if $T^n$ has a symmetry that reflects "orthogonally to $v$". The condition (S1) implies that length minimizers in $\mathcal{C}_v$ are (closed) minimal geodesics.

The main results of Bolotin and Rabinowitz [10] yield homoclinic and heteroclinic minimal geodesics $\gamma$. All these geodesics $\gamma$ have $A_{\mathrm{mix}}(\gamma) = \{v\}$, thus they are homologically homoclinic in the sense of Subsection 5.5, and $t \mapsto \gamma(t)$ is asymptotic to some $\tau_+$ for $t \to +\infty$ and asymptotic to some $\tau_-$ for $t \to -\infty$, where $\tau_\pm$ are length minimizers in $\mathcal{C}_v$. In the case $p = 1$, i. e., if there is a unique closed minimal geodesic $\tau_1$ in $\mathcal{C}_v$, then there exist at least $n$ homoclinic minimal geodesics [10, Theorem 1.7], all of them are asymptotic to $\tau_1$ for $t \to +\infty$ and $t \to -\infty$. For $p > 1$, there exist at least $n + p - 1$ heteroclinic minimal geodesic, with the above properties [10, Theorem 1.8]

All these minimal geodesics are homologically homoclinic, thus they are of a different type than the ones detected by our methods. This fact sheds some light on the minimal number $\mathcal{N}_{\min}(T^3)$ of minimal geodesics on $T^3$.

To be more precise, let $\mathcal{N}(T^3, g)$ be the number of geometrically distinct minimal geodesics on $(T^3, g)$, and define

$$\mathcal{N}_{\min}(T^3) := \min\bigl\{\mathcal{N}(T^n, g) \mid g \text{ is a Riemannian metric on } T^3\bigr\}.$$

We stated in Subsection 1.4 that we expect $15 \le \mathcal{N}_{\min}(T^3) \le 27$, and it would be interesting to determine the precise value of $\mathcal{N}_{\min}(T^3)$.

Suppose that $g$ is a Hedlund metric on $T^n$ with finitely many minimal geodesics. Let us assume that $g$ also satisfies (S1) and (S2) [10, Theorem 1.8] – all explicit constructions of Hedlund metrics known to the authors do so.

Then the results by Bangert [6] and Bolotin–Rabinowitz [10] provide at least 9 homologically homoclinic minimal geodesics and our method yields at least 12 homologically non-homoclinic minimal geodesics. Thus one has $\mathcal{N}(T^3, g) \geq 21$ for all Hedlund metrics $g$ satisfying the assumptions of Bolotin and Rabinowitz [10].

As a conclusion we see that $\mathcal{N}_{\min}(T^3)$ is either essentially larger than 15, or one has to use much finer construction methods in order to obtain a metric $g$ with $\mathcal{N}(T^3, g)$ close to 15. It thus would be desirable to find a refined existence theorem for homologically homoclinic minimal geodesics in the style of Bolotin–Rabinowitz [10] without assumptions (S1) and (S2), strengthening Bangert's existence result [6, Theorem 4.4]. Such an improved existence result would imply that $\mathcal{N}_{\min}(T^3)$ is essentially larger than 15.

In summary, we see that our lower bound on the number of homologically non-homoclinic minimal geodesics is optimal on $T^3$, but it is still unknown whether the existing lower bound on the number of homologically homoclinic minimal geodesics is optimal.

## Appendix A. Length and minimizing geodesics

We collect some facts about the length functional and on distance minimizing geodesics that are well-known or straightforward consequences of well-known facts.

A.1. **Length of curves.** Let $(X, d)$ be a metric space. For a continuous curve $\gamma\colon I \to X$ on a non-empty interval $I \subset \mathbb{R}$, we define the *length* by

$$\mathcal{L}(\gamma) \coloneqq \sup\Big\{\sum_{i=1}^{k} d(\gamma(t_i), \gamma(t_{i-1})) \;\Big|\; k \in \mathbb{N}_0,\ t_0, \dots, t_k \in I,\ t_0 < t_1 \dots < t_k\Big\} \in [0, \infty]\,.$$

If $(X, d)$ arises from a Riemannian manifold and if $\gamma$ is piecewise $C^1$, then $\mathcal{L}(\gamma) = \int_I \|\dot\gamma(t)\|\,\mathrm{d}t \in [0, \infty]$.

A.2. **Limiting geodesics.** We provide a proof of Lemma 2.3:

*Proof of Lemma 2.3.* It follows from the theorem of Picard and Lindelöf for suitable ordinary differential equations, in particular from the smooth dependence of the solution on the initial conditions, that $\sigma_i$ converges to $\sigma_\infty$, in the sense of uniform convergence on all compact intervals.

The induced distance function $\hat d\colon \widehat M \times \widehat M \to \mathbb{R}_{\geq 0}$ is continuous; thus, as all curves $\sigma_i$ are minimizing, this allows the following calculation for all $s, t \in \mathbb{R}$ with $s < t$:

$$\begin{aligned} \hat d\big(\sigma_\infty(t), \sigma_\infty(s)\big) &= \hat d\big(\lim_{i\to\infty} \sigma_i(t), \lim_{i\to\infty} \sigma_i(s)\big) \\ &= \lim_{i\to\infty} \underbrace{\hat d\big(\sigma_i(t), \sigma_i(s)\big)}_{=t-s} = t - s\,. \end{aligned}$$

Therefore, $\sigma_\infty$ is minimizing as well. ☐

A.3. **Existence of minimal geodesics and the fundamental group.**

**Proposition A.1.** *Let $(M, g)$ be a closed connected Riemannian manifold. Then there exists a minimal geodesic on $M$ if and only if $\pi_1(M)$ is infinite.*

*Proof.* The universal covering $\widetilde M$ of $M$ has infinite diameter $\operatorname{diam} \widetilde M$ if and only if $\pi_1(M)$ is infinite. If there exists a minimal geodesic on $M$, then clearly, $\operatorname{diam} \widetilde M = \infty$.

Conversely, let $\operatorname{diam} \widetilde M = \infty$. Then there are points $p_i, q_i \in \widetilde M$ with $\ell_i \coloneqq \widetilde d(p_i, q_i)/2 \to \infty$. We choose a minimizing geodesic $\sigma_i\colon [-\ell_i, \ell_i] \to \widetilde M$ with $\sigma_i(-\ell_i) =$

$p_i$ and $\sigma_i(\ell_i) = q_i$, parametrized by arclength. As the deck transformation group acts cocompactly on $\widetilde{M}$ and on the unit tangent bundle of $\widetilde{M}$, we can assume – without loss of generality – that there is a compact set $K$ containing $\dot{\sigma}_i(0)$ for all $i \in \mathbb{N}$. This allows us to pass to a subsequence such that $v_\infty := \lim_{i\to\infty} \dot{\sigma}_i(0)$ exists. We apply Lemma 2.3 for $a_i := -\ell_i$ and $b_i := \ell_i$. The resulting geodesic $\sigma_\infty \colon \mathbb{R} \to \widetilde{M},\ t \mapsto \exp(tv_\infty)$ is globally distance minimizing, and thus – by definition – its projection to $M$ is a minimal geodesic. □

A.4. **Length-minimizing curves.** Next, we prove that the infimum in Definition 3.5 is attained and that it does not matter whether the infimum ranges over curves of high or low regularity, as long as they are continuous. Let us recall a lemma that is a reformulation of a lemma in Sakai's book.

**Lemma A.2** ([32, Lemma V.1.5 (1)]). *Let $(M, g)$ be a closed connected Riemannian manifold. Then, in every nontrivial free homotopy class of closed curves $S^1 \to M$, there exists a curve of minimal length. This curve can be parametrized by arclength and then it is a closed geodesic.*

In the following, if $\gamma$ is a closed curve in $M$, then $h(\gamma)$ denotes the element in $H_1(M;\mathbb{Z})_\mathbb{R}$ represented by $\gamma$ via the Hurewicz homomorphism.

**Lemma A.3.** *Let $M$ be a closed connected Riemannian manifold and let $x \in H_1(M;\mathbb{Z})_\mathbb{R} \setminus \{0\}$. Then there is a closed geodesic $\gamma : [0, N(x)] \to M$ with $h(\gamma) = x$ and*

$$\mathcal{L}(\gamma) = \inf\bigl\{\mathcal{L}(\tau) \bigm| \tau \text{ is a continuous loop with } h(\tau) = x\bigr\}\,.$$

The lemma can be proved via arguments analogous to a related proof in Sakai's book [32, Lemma V.1.5 (2)], enriched with standard smoothing techniques [26, §16 and 17].

A.5. **Distance estimates in finite-sheeted coverings.**

**Lemma A.4.** *Let $Q_1 \xrightarrow{p} Q_2 \xrightarrow{\pi} M$ be a sequence of Riemannian coverings, where $M$ is a closed Riemannian manifold with metric $g$. (Then also $Q_1 \xrightarrow{\pi\circ p} M$ is a covering.) We assume that the coverings $p$, $\pi$, and $\pi\circ p$ are normal, i. e., the deck transformation groups act transitively (and freely) on the fibres of the respective coverings. Let $g_i$ be the pullback of $g$ to $Q_i$ and let $d_i$ be the induced distance function on $Q_i$.*

(1) *Then, for all $x, y \in Q_1$ we have*
$$d_2\bigl(p(x), p(y)\bigr) \le d_1(x, y)\,.$$
(2) *If $Q_1 \xrightarrow{p} Q_2$ is a finite covering, then there is a $C < \infty$, such that for all $x, y \in Q_1$:*
$$d_1(x, y) \le d_2\bigl(p(x), p(y)\bigr) + C$$

*Proof.* The first part is clear because $p$ is a local isometry. We show the second part: For $\hat{x} \in Q_2$ we define

$$m(\hat{x}) := \operatorname{diam}_{d_1} p^{-1}(\hat{x}) = \max\left\{d_1(x, y) \mid x, y \in p^{-1}(\hat{x})\right\}\,.$$

It is easy to see that $m \colon Q_2 \to \mathbb{R}$ is a continuous function. We show that $m$ factors over $\pi \colon Q_2 \to M$: As $\pi$ is normal, for all $\hat{x}_1, \hat{x}_2 \in Q_2$ with $\pi(\hat{x}_1) = \pi(\hat{x}_2)$, there exists a deck transformation $\hat{f}$ of $Q_2 \to M$ with $\hat{f}(\hat{x}_1) = \hat{x}_2$. Because $\pi\circ p$ is normal, $\hat{f}$ lifts to a deck transformation $f$ of $p$; in particular, $\hat{f}$ maps $p^{-1}(\hat{x}_1)$ bijectively to $p^{-1}(\hat{x}_2)$. Thus $m$ is invariant under the deck transformations of $Q_2 \to M$. Therefore, $m$ factors as $m = m' \circ \pi$ for some continuous map $m' \colon M \to \mathbb{R}$. In particular, the function $m$ attains its maximum and we define

$$C \coloneqq \max_{\hat{x} \in Q_2} m(\hat{x}) .$$

Let $x, y \in Q_1$. The Riemannian manifold $Q_2$ is complete and so by the Hopf–Rinow theorem there is a path $\gamma$ in $Q_2$ from $p(x)$ to $p(y)$ of length $\ell \coloneqq d_2\bigl(p(x), p(y)\bigr)$. Let $\gamma'$ be the lift of $\gamma$ to $Q_1$ starting in $x$. If $y'$ is the endpoint of $\gamma'$, then $p(y')$ is the endpoint of $\gamma$ and so $p(y') = p(y)$. By construction of $C$, we have $d_1\bigl(y', y\bigr) \leq C$. Furthermore, the existence of $\gamma'$, which has length $\ell$, shows $d_1\bigl(x, y'\bigr) \leq \ell$. Then the statement follows from the triangle inequality. $\square$

## Appendix B. Other aspects of the stable norm

We recall (special cases of) Federer's notion of mass $\|\bullet\|_g^K$, which is a group norm on $H_1(M; K)$ for $K \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. We then show that the norms $\|\bullet\|_g^{\mathbb{R}}$ and $\|\bullet\|_{\mathrm{st}}$ are equal.

This equivalence is well-known, and – as already mentioned in Section 3.3 – sometimes attributed to Gromov's book on metric structures on Riemannian manifolds [18] or its French version [16]. The equivalence is based on work by Federer, requiring involved notation. We include a proof to keep our presentation self-contained.

**Definition B.1** (mass of 1-homology classes)**.** Let $(M, g)$ be a Riemannian manifold and $K \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. For a singular 1-chain $a = \sum_{i=1}^m a_i \sigma_i \in C_1(M; K)$ with coefficients $a_i$ in $K$, we define

$$\|a\|_g \coloneqq \sum_{i=1}^m |a_i| \cdot \mathcal{L}(\sigma_i) .$$

For $\alpha \in H_1(M; K)$, we define the *$K$-mass* of $\alpha$ as

$$\|\alpha\|_g^K \coloneqq \inf_{a \in \alpha} \|a\|_g \; .$$

It follows directly from the definition that the map $\|\bullet\|_g^K : H_1(M; K) \to \mathbb{R}$ is symmetric and satisfies the triangle inequality; i.e., it is what is sometimes referred to as a *group semi-norm*. The manifold $\mathbb{R}^2 \setminus \{0\}$ with the Euclidean metric shows that in general we do not get positive definiteness of $\|A\|_g^K$ without further assumptions. In the following, we therefore often require that the manifold is closed.

The group semi-norm $\|\bullet\|_g^{\mathbb{R}}$ is $\mathbb{R}$-homogeneous and thus a semi-norm on a vector space in the usual sense.

**Lemma B.2.** *Let $(M, g)$ be a Riemannian manifold. Then, the change-of-coefficients map $\bullet_{\mathbb{R}} \colon H_1(M; \mathbb{Q}) \to H_1(M; \mathbb{R})$ is isometric: For all $\alpha \in H_1(M; \mathbb{Q})$, we have*

$$\|\alpha_{\mathbb{R}}\|_g^{\mathbb{R}} = \|\alpha\|_g^{\mathbb{Q}} \; .$$

*Proof.* The change-of-coefficients map is induced by the inclusion $i \colon C_*(M; \mathbb{Q}) \to C_*(M; \mathbb{R})$ of the singular chain complexes. The map $i$ is isometric in each degree and has dense image with respect to $\|\bullet\|_g$. Approximating boundaries therefore shows that $\arg_{\mathbb{R}} = H_1(i) \colon H_1(M; \mathbb{Q}) \to H_1(M; \mathbb{R})$ is isometric with respect to the induced semi-norms $\|\cdot\|_g^{\mathbb{Q}}$ and $\|\cdot\|_g^{\mathbb{R}}$ on homology [33, Lemma 2.9] [22, Lemma 1.7] (the cited proofs carry over to this slightly more general setting). $\square$

**Lemma B.3.** *Let $(M, g)$ be a Riemannian manifold. Then $\|\cdot\|_g^{\mathbb{R}}$ on $H_1(M;\mathbb{R})$ extends the homogenisation (in the sense of Proposition 3.2) of $\|\cdot\|_g^{\mathbb{Z}}$.*

*Proof.* Let $\alpha \in H_1(M;\mathbb{Z})$ and we denote its image in $H_1(M;\mathbb{R})$ with respect to the change-of-coefficients map $H_1(M;\mathbb{Z}) \to H_1(M;\mathbb{R})$ by $\alpha_{\mathbb{R}}$ (and similarly for $\mathbb{Q}$). Using Lemma B.2 and the definition of $\|\cdot\|_g^{\mathbb{Q}}$, we obtain

$$\|\alpha_{\mathbb{R}}\|_g^{\mathbb{R}} = \|\alpha_{\mathbb{Q}}\|_g^{\mathbb{Q}} = \inf_{k\in\mathbb{N}} \frac{1}{k} \cdot \|k \cdot \alpha\|_g^{\mathbb{Z}} .$$

According to Proposition 3.2, the right-hand side equals to the homogenisation of $\|\cdot\|_g^{\mathbb{Z}}$ on $\alpha$. $\square$

**Proposition B.4** (mass and stable norm)**.** *Let $(M, g)$ be a closed connected Riemannian manifold. Then $\|\cdot\|_g^{\mathbb{R}}$ and $\|\cdot\|_{\mathrm{st}}$ coincide on $H_1(M;\mathbb{R})$.*

*Proof.* We first show $\|\cdot\|_g^{\mathbb{R}} \leq \|\cdot\|_{\mathrm{st}}$: Let $\alpha \in H_1(M;\mathbb{Z})_{\mathbb{R}}$. Then there is a loop $\gamma$ of length $N(\alpha)$ representing $\alpha$. If we re-interpret $\gamma$ as a 1-cycle, we get

$$\|\alpha_{\mathbb{R}}\|_g^{\mathbb{R}} \leq \|\alpha\|_g^{\mathbb{Z}} \leq \|\gamma\|_g = \mathcal{L}(\gamma) = N(\gamma) .$$

Applying homogenisation (Proposition 3.2) and using the fact that $\|\cdot\|_g^{\mathbb{R}}$ is homogeneous, we obtain $\|\cdot\|_g^{\mathbb{R}} \leq \|\cdot\|_{\mathrm{st}}$.

We now prove the converse inequality. As rational classes are dense in $H_1(M;\mathbb{R})$ with respect to $\|\cdot\|_g^{\mathbb{R}}$ and $\|\cdot\|_{\mathrm{st}}$, it suffices to prove the estimate $\|\cdot\|_g^{\mathbb{R}} \geq \|\cdot\|_{\mathrm{st}}$ on rational classes. Moreover, in view of Lemma B.2, we may replace $\|\cdot\|_g^{\mathbb{R}}$ with $\|\cdot\|_g^{\mathbb{Q}}$.

Let $\alpha \in H_1(M;\mathbb{Q})$ and $\varepsilon \in \mathbb{R}_{>0}$. Then, there is a singular chain $a = \sum_{i=1}^{m} a_i\sigma_i \in C_1(M;\mathbb{Q})$ representing $\alpha$ with $\|a\|_g \leq \|\alpha\|_g^{\mathbb{Q}} + \varepsilon$. We will replace $a$ with loops that efficiently represent multiples of $\alpha$. More precisely, let $k_0 \in \mathbb{N}$ be a common multiple of the denominators appearing in $a_1, \dots, a_m \in \mathbb{Q}$. There exists a chain $b = \sum_{i=1}^{m'} \sigma_i' \in C_1(M;\mathbb{Z})$, consisting of a sum of *loops* $\sigma_i'$, with the following properties:

$$[b] = k_0 \cdot \alpha \quad \text{and} \quad \|b\|_g = \sum_{i=1}^{m'} \mathcal{L}(\sigma_i') \leq k_0 \cdot \big(\|\alpha\|_g^{\mathbb{Q}} + \varepsilon\big)$$

Indeed, because $k_0 \cdot a$ is an integral cycle, we can re-organise the singular 1-simplices appearing in $k_0 \cdot a$ (with multiplicities/orientations given by their coefficients) into a sum of loops. This procedure does not affect the homology class.

We parametrize the singular simplices $\sigma_i'$ on $[0, \ell_i]$; moreover, for each $i$, we choose a path $\tau_i$ from $\sigma_i(0)$ to $\sigma_{i+1}(0)$ with $\mathcal{L}(\tau_i) \leq \operatorname{diam} M$. For $k_1 \in \mathbb{N}$, let us consider

$$w_{k_1} \coloneqq \underbrace{\sigma_1' * \cdots * \sigma_1'}_{k_1\text{-times}} * \tau_1 * \underbrace{\sigma_2' * \cdots * \sigma_2'}_{k_1\text{-times}} * \tau_2 * \cdots * \underbrace{\sigma_{m'}' * \cdots * \sigma_{m'}'}_{k_1\text{-times}} * \bar{\tau}_{m'-1} * \cdots * \bar{\tau}_1,$$

which provides a loop based at $\sigma_1'(0)$ with $[w_{k_1}] = k_1 \cdot [k_0 \cdot a] \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ and of length

$$\mathcal{L}(w_{k_1}) \leq k_1 \cdot \sum_{i=1}^{m'} \mathcal{L}(\sigma_i') + 2(m' - 1) \cdot \operatorname{diam} M .$$

Thus

$$\frac{1}{k_1} \cdot N(k_1 \cdot [k_0 \cdot a]) \leq \frac{1}{k_1} \cdot \mathcal{L}(w_{k_1}) \leq \sum_{i=1}^{m'} \mathcal{L}(\sigma'_i) + \frac{2(m'-1)}{k_1} \cdot \operatorname{diam} M \,.$$

As $\|\bullet\|_{\mathrm{st}}$ was defined as the homogenisation of $N(\bullet)$, the left-hand side converges for $k_1 \to \infty$ to $\big\|[k_0 \cdot a]\big\|_{\mathrm{st}} = k_0 \cdot \|\alpha\|_{\mathrm{st}}$. On the other hand, the limit of the right-hand side is bounded from above by $k_0 \cdot (\|\alpha\|_g^{\mathbb{Q}} + \varepsilon)$. Dividing by $k_0$ and taking $\varepsilon \to 0$, we get $\|\alpha\|_{\mathrm{st}} \leq \|\alpha\|_g^{\mathbb{R}}$. □

If $M$ is closed, for $\omega \in \Omega^1(M)$, we define the supremum norm (or $L^\infty$-norm or $C^0$-norm) as

$$\|\omega\|_{L^\infty} \coloneqq \sup_{x \in M} \big|\omega(x)\big|_g \,.$$

For 1-forms, this coincides with Gromov's notion of *comass* [18, 4.15].

**Proposition B.5** (duality and the comass)**.** *Let $M$ be a closed connected Riemannian manifold. Then, for all $\varphi \in H^1_{\mathrm{dR}}(M)$ we have:*

$$\|\varphi\|^*_{\mathrm{st}} = \inf_{\omega \in \varphi} \|\omega\|_{L^\infty} \ .$$

*Proof of the inequality $\leq$.* Let $\omega \in \Omega^1(M)$. Take $v \in H_1(M;\mathbb{Z})_{\mathbb{R}}$ and a closed geodesic $\gamma$ representing $v$ with $\mathcal{L}(\gamma) = N(v)$. Then

$$\big|\langle[\omega], v\rangle\big| \quad = \quad \Big|\int_\gamma \omega\Big| \ \leq \ \mathcal{L}(\gamma) \cdot \|\omega\|_{L^\infty} \ = \ N(v) \cdot \|\omega\|_{L^\infty} \ .$$

We apply homogenisation. As the left-hand side is homogeneous in $v$, we obtain

$$\big|\langle[\omega], v\rangle\big| \quad \leq \quad \|v\|_{\mathrm{st}} \cdot \|\omega\|_{L^\infty} \ . \tag{B.1}$$

By homogeneity of this inequality, this estimate (B.1) also holds for all rational classes $v \in H_1(M;\mathbb{Q})$; and then by density for all $v \in H_1(M;\mathbb{R})$. Using the defining Equation (3.1) of the dual of the stable norm, this implies $\|[\omega]\|^*_{\mathrm{st}} \leq \|\omega\|_{L^\infty}$, which proves the claimed inequality. □

The proof of the inequality "$\geq$" requires the construction of suitable 1-forms $\omega$. A proof is given by Gromov [18, 4.35].

**Comparison to the literature.** Gromov [16, Proposition 2.22] calls the vector space norm obtained from a homogenisation of a group norm $\|\bullet\|$ in the sense of Proposition 3.2 *norme limite* and denotes it as $\|\bullet\|^{lim}$. In the English edition [18, 4.17], the norm is denoted in the same way and called *limit norm*, but the "Proposition 2.22" to which he refers differs in the English version.

Gromov uses the symbol $\|\bullet\|$ or $\|\bullet\|_{H_1}$ for all the norms $\|\bullet\|_g : C_1(M;K) \to K$, $\|\bullet\|_g^{\mathbb{Z}} : H_1(M;\mathbb{Z}) \to \mathbb{Z}$, and $\|\bullet\|_g^{\mathbb{R}} : H_1(M;\mathbb{R}) \to \mathbb{R}$. This does not lead to ambiguities as the argument $\alpha$ of $\|\alpha\|$ will determine, which (group) norm is meant.

## Appendix C. More on minimal geodesics in dimension 2

In this appendix let $M$ be an orientable, closed, connected surface of genus at least 1 with Riemannian metric $g$. For $\alpha \in \pi_1(M)$ let $[\alpha]_{\mathrm{free}}$ be its free homotopy class. The length functional attains its minimum in any free homotopy class $[\alpha]_{\mathrm{free}}$, and we write $N_0\big([\alpha]_{\mathrm{free}}\big)$ for the minimum of this length. Let us recall the following classical lemma, for which one may find a proof, for example, in an article by Bleecker [8].

**Lemma C.1.** *Let $M$ be an orientable, closed, connected surface of genus at least 1 with Riemannian metric $g$. Let $\gamma\colon [0,L] \to M$ be a closed geodesic in $M$ minimizing length within its free homotopy class $[\alpha^k]_{\mathrm{free}}$ where $\alpha \in \pi_1(M)$ is primitive and $k \in \mathbb{N}$. Then $\gamma\colon [0,L/k] \to M$ is a closed geodesic in $M$ minimizing length within its free homotopy class $[\alpha]_{\mathrm{free}}$, and $\gamma$ extends to a periodic curve $\mathbb{R} \to M$ of period $L/k$. in particular, we have for every $\beta \in \pi_1(M)$:*

$$N_0\big(\big[\beta^k\big]_{\mathrm{free}}\big) = k \cdot N_0\big([\beta]_{\mathrm{free}}\big)\,.$$

*Proof.* We choose $x_0 \coloneqq \gamma(0)$ and a lift $\widetilde{x}_0 \in \widetilde{M}$ as basepoints. This allows to identify $\pi_1(M,x_0)$ with the deck transformations of $\widetilde{M} \to M$. We consider $\widehat{M} \coloneqq \widetilde{M}/\langle\alpha^k\rangle$, which is diffemorphic to $S^1\times\mathbb{R}$. Then $\gamma$ lifts to a simple closed geodesic $\widehat{\gamma}\colon [0,L] \to \widehat{M}$ with $\widehat{\gamma}(0) = [\widetilde{x}_0] =: \widehat{x}_0$. The closed geodesic $\widehat{\gamma}$ generates the infinite cyclic group $\pi_1(\widehat{M}) \cong H_1(\widehat{M};\mathbb{Z})$. The group $\{\alpha^m \mid m \in \mathbb{Z}\} \subset \pi_1(M,x_0)$ acts freely and isometrically on $\widehat{M}$, and $\alpha^k$ acts trivially.

We claim that the curves $\widehat{\gamma}$ and $\alpha \cdot \widehat{\gamma}$ necessarily have to intersect in at least two points. The Jordan curve theorem implies that $\widehat{M} \setminus \mathrm{image}(\widehat{\gamma})$ has two components, and orientation allows us to call them the right and the left connected component of $\widehat{\gamma}$ (this requires the choice of some convention, but any choice will be fine). If $\widehat{\gamma}$ and $\alpha\cdot\widehat{\gamma}$ do not intersect, then $\alpha\cdot\widehat{\gamma}$ has to run entirely in the left (or right) connected component of $\widehat{\gamma}$. As $\alpha$ preserves orientation, $\alpha^2 \cdot \widehat{\gamma}$ has to run entirely in the left (or right) connected component of $\alpha \cdot \widehat{\gamma}$ and thus in the left (or right) connected component of $\widehat{\gamma}$. By induction, we get the same statement for $\alpha^k \cdot \widehat{\gamma}$ instead of $\alpha^2 \cdot \widehat{\gamma}$, but this is obviously in contradiction to $\alpha^k \cdot \widehat{\gamma} = \widehat{\gamma}$.

If the curves $\widehat{\gamma}$ and $\alpha \cdot \widehat{\gamma}$ intersect in precisely one point, then the Jordan curve theorem requires that away from this point $\widehat{\gamma}$ is on one side of $\alpha \cdot \widehat{\gamma}$ and thus a similar argument yields a contradiction. The claim is thus proven.

We now prove that $\widehat{\gamma}$ and $\alpha \cdot \widehat{\gamma}$ coincide (up to reparametrization): If they do not coincide, we can use cut and paste constructions at the intersection points to produce a loop in $\hat{M}$ shorter than $L$, freely homotopic to $\widehat{\gamma}$, which is again a contradiction.

Thus, $\hat{\gamma}$ is invariant under the action of $\alpha$ (up to shift in the parameter) and all statements of the lemma follow directly from this. □

**Proposition C.2.** *Let $M$ be an orientable, closed, connected surface of genus at least 1 with Riemannian metric $g$. Let $\gamma\colon \mathbb{R} \to M$ be a (non-constant) closed geodesic with respect to $g$, that minimizes length in its free homotopy class. Then, $\gamma$ is a minimal geodesic.*

*Proof.* As always, we assume that the geodesic $\gamma$ is parametrized by arclength. Let $\ell > 0$ be the period of $\gamma$. As $\gamma$ is not constant, the free homotopy class of $\gamma\big|_{[0,\ell]}$ is non-trivial and thus also the class $\alpha \coloneqq [\gamma\big|_{[0,\ell]}] \in \pi_1\big(M,\gamma(0)\big)$ is non-trivial.

Let $\widetilde{\gamma}\colon \mathbb{R} \to \widetilde{M}$ be a lift of $\gamma$. Because the fundamental group of the surface $M$ is torsion-free, $\alpha$ has infinite order. Moreover, $\alpha^k \cdot \widetilde{\gamma}(0)$ converges for $k \to \infty$ to a point $p_\infty \in \partial\widetilde{M}$, where $\partial\widetilde{M}$ is the boundary at infinity of $\widetilde{M}$. Similarly, the sequence converges to a point $p_{-\infty} \in \partial\widetilde{M}$ for $k \to -\infty$.

Assume for a contradiction that $\widetilde{\gamma}$ is not a line. Then there exist $t_0, t_1 \in \mathbb{R}$ with $t_0 < t_1$ and

$$\widetilde{d}\big(\widetilde{\gamma}(t_0), \widetilde{\gamma}(t_1)\big) < t_1 - t_0 = \mathcal{L}\big(\gamma|_{[t_0,t_1]}\big)\,.$$

By increasing $t_1$ and decreasing $t_0$ we may assume that $t_i = k_i\ell$ with $k_i \in \mathbb{Z}$ for $i \in \{0,1\}$, and $k \coloneqq k_1 - k_0 \geq 2$. By shifting in the domain by $t_0$ and in the range by $\alpha^{k_0}$, we may assume $t_0 = 0$ and $t_1 = k\ell$ with $k \in \mathbb{N}_{\geq 2}$. There is a curve $\widetilde{\tau}$ of length $L$ less than $k\ell$ from $\widetilde{\gamma}(0)$ to $\widetilde{\gamma}(k\ell)$. The projections of $\widetilde{\gamma}|_{[0,k\ell]}$ and $\widetilde{\tau}$ to $M$ are loops and will be denoted as $\gamma'$ and $\tau$. Both of them represent the free homotopy class defined by $\alpha^k$. As $\tau$ is shorter than $\gamma'$, we have $N_0\big([\alpha^k]_{\mathrm{free}}\big) \leq \mathcal{L}(\tau) < k\ell$. This contradicts Lemma C.1 and our proposition is shown. $\square$

Note that orientability is crucial for these results. For example, if $M$ is the Klein bottle, one can construct a Riemannian metric $g$ on $M$ such that Lemma C.1 and Proposition C.2 do no longer hold for this non-orientable surface $M$ and the metric $g$.

## References


[1] Ammann, B. Minimale Geodätische auf Mannigfaltigkeiten mit nilpotenter Fundamentalgruppe. Diplomarbeit, Universität Freiburg, Germany, 1994. DOI: 10.5283/epub.53183. Cited on page: 1, 5, 6, 15, 16, 27, 29, 44, 45, 47

[2] Ammann, B. Minimal geodesics and nilpotent fundamental groups. *Geom. Dedicata 67* (1997), 129–148. Cited on page: 5, 6, 29, 44, 45

[3] Ammann, B., and Löh, C. Edges of symmetric polytopes. Work in progress. Cited on page: 2

[4] Babenko, I., and Balacheff, F. Sur la forme de la boule unité de la norme stable unidimensionnelle. *Manuscripta Math. 119*, 3 (2006), 347–358. Cited on page: 16, 44

[5] Bangert, V. Mather sets for twist maps and geodesics on tori. In *Dynamics reported, Vol. 1*, vol. 1 of *Dynam. Report. Ser. Dynam. Systems Appl.* Wiley, Chichester, 1988, pp. 1–56. Cited on page: 43

[6] Bangert, V. Minimal geodesics. *Ergod. Th. & Dynam. Sys. 10*, 2 (1990), 263–286. Cited on page: 1, 2, 3, 5, 6, 10, 11, 13, 14, 15, 17, 23, 27, 28, 29, 40, 43, 44, 49, 50

[7] Bangert, V. Geodesic rays, Busemann functions and monotone twist maps. *Calc. Var. Partial Differential Equations 2*, 1 (1994), 49–63. Cited on page: 16

[8] Bleecker, D. D. The Gauss-Bonnet inequality and almost-geodesic loops. *Advances in Math. 14* (1974), 183–193. Cited on page: 55

[9] Bliss, G. A. The geodesic lines on the anchor ring. *Ann. of Math. (2) 4*, 1 (1902), 1–21. Cited on page: 43

[10] Bolotin, S. V., and Rabinowitz, P. H. Minimal heteroclinic geodesics for the $n$-torus. *Calc. Var. Partial Differential Equations 9*, 2 (1999), 125–139. Cited on page: 49, 50

[11] Burago, D. Periodic metrics. *Advances in Soviet Mathematics 9* (1992), 205–210. Cited on page: 6, 18, 22

[12] Burago, D., Burago, Y., and Ivanov, S. *A course in metric geometry*, vol. 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Cited on page: 3, 6, 12

[13] Burago, D., Ivanov, S., and Kleiner, B. On the structure of the stable norm of periodic metrics. *Math. Res. Lett. 4*, 6 (1997), 791–808. Cited on page: 16

[14] Federer, H. Real flat chains, cochains and variational problems. *Indiana Univ. Math. J. 24* (1974/75), 351–407. Cited on page: 14

[15] Freedman, M., Hass, J., and Scott, P. Closed geodesics on surfaces. *Bull. London Math. Soc. 14*, 5 (1982), 385–391. Cited on page: 44

[16] Gromov, M. *Structures métriques pour les variétés riemanniennes*, vol. 1 of *Textes Mathématiques [Mathematical Texts]*. CEDIC, Paris, 1981. Cited on page: 52, 54

[17] Gromov, M. Hyperbolic groups. In *Essays in group theory*, vol. 8 of *Math. Sci. Res. Inst. Publ.* Springer, New York, 1987, pp. 75–263. Cited on page: 1, 44

[18] Gromov, M. *Metric structures for Riemannian and non-Riemannian spaces*. Birkhäuser Boston Inc., Boston, MA, 1999. Cited on page: 10, 11, 14, 18, 52, 54

[19] Hedlund, G. A. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. of Math. (2) 33*, 4 (1932), 719–739. Cited on page: 1, 3, 5, 6, 29, 42, 44

[20] Jotz, M. Hedlund metrics and the stable norm. *Differential Geom. Appl. 27*, 4 (2009), 543–550. Cited on page: 16, 44

[21] Klingenberg, W. Geodätischer Fluss auf Mannigfaltigkeiten vom hyperbolischen Typ. *Invent. Math. 14* (1971), 63–82. Cited on page: 1, 44

[22] Löh, C. *$\ell^1$-Homology and Simplicial Volume*. PhD thesis, WWU Münster, 2007. `http://nbn-resolving.de/urn:nbn:de:hbz:6-37549578216`. Cited on page: 53

[23] Massart, D. Normes stables des surfaces. *C. R. Acad. Sci. Paris Sér. I Math. 324*, 2 (1997), 221–224. Cited on page: 44

[24] Massart, D. Stable norms of surfaces: local structure of the unit ball of rational directions. *Geom. Funct. Anal. 7*, 6 (1997), 996–1010. Cited on page: 44

[25] Meeks, III, W. H., and Patrusky, J. Representing codimension-one homology classes by embedded submanifolds. *Pacific J. Math. 68*, 1 (1977), 175–176. Cited on page: 42

[26] Milnor, J. *Morse theory*, vol. No. 51 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1963. Based on lecture notes by M. Spivak and R. Wells. Cited on page: 13, 51

[27] Montealegre, P. On the stable norm of slit tori and the Farey sequence, 2023. Preprint, arXiv: 2310.05570. Cited on page: 44

[28] Morse, H. M. A fundamental class of geodesics on any closed surface of genus greater than one. *Trans. Amer. Math. Soc. 26*, 1 (1924), 25–60. Cited on page: 1, 42, 44

[29] Navas, A. *Groups of circle diffeomorphisms*, spanish ed. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 2011. Cited on page: 43

[30] Novik, I. A tale of centrally symmetric polytopes and spheres. In *Recent trends in algebraic combinatorics*, vol. 16 of *Assoc. Women Math. Ser.* Springer, 2019, pp. 305–331. Cited on page: 39

[31] Pólya, G., and Szegő, G. *Problems and theorems in analysis. I.* Classics in Mathematics. Springer, 1998. Series, integral calculus, theory of functions, Translated from the German by Dorothee Aeppli, Reprint of the 1978 English translation. Cited on page: 12

[32] Sakai, T. *Riemannian geometry*, vol. 149 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1996. Translated from the 1992 Japanese original by the author. Cited on page: 51

[33] Schmidt, M. *$L^2$-Betti Numbers of $\mathcal{R}$-Spaces and the Integral Foliated Simplicial Volume*. PhD thesis, Westfälische Wilhelms-Universität Münster, 2005. http://nbn-resolving.de/urn:nbn:de:hbz:6-05699458563. Cited on page: 53

[34] Straszewicz, S. Über exponierte Punkte abgeschlossener Punktemengen. *Fund. Math. 24* (1935), 139–143. Cited on page: 40



B. Ammann, Fakultät für Mathematik, Universität Regensburg, Regensburg, Germany
*Email address*: bernd.ammann@mathematik.uni-regensburg.de

C. Löh, Fakultät für Mathematik, Universität Regensburg, Regensburg, Germany
*Email address*: clara.loeh@mathematik.uni-regensburg.de