

Counting Periodic Trajectories of Finsler Billiards

Paule V.M. BLAGOJEVIĆ^{†1†2}, Michael HARRISON^{†3}, Serge TABACHNIKOV^{†4}
and Günter M. ZIEGLER^{†1}

^{†1} *Institut für Mathematik, FU Berlin, Arnimallee 2, 14195 Berlin, Germany*

E-mail: blagojevic@math.fu-berlin.de, ziegler@math.fu-berlin.de

URL: <http://page.mi.fu-berlin.de/blagojevic/>

<http://www.mi.fu-berlin.de/math/groups/discgeom/ziegler>

^{†2} *Mathematical Institut SASA, Knez Mihailova 36, 11000 Beograd, Serbia*

^{†3} *Department of Mathematical Science, Carnegie Mellon University,*

Pittsburgh, PA 15213, USA

E-mail: mah5044@gmail.com

URL: <https://sites.google.com/site/michaelharrisonmath/>

^{†4} *Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA*

E-mail: tabachni@math.psu.edu

URL: <http://www.personal.psu.edu/sot2/>

Received September 11, 2019, in final form March 25, 2020; Published online April 03, 2020

<https://doi.org/10.3842/SIGMA.2020.022>

Abstract. We provide lower bounds on the number of periodic Finsler billiard trajectories inside a quadratically convex smooth closed hypersurface M in a d -dimensional Finsler space with possibly irreversible Finsler metric. An example of such a system is a billiard in a sufficiently weak magnetic field. The r -periodic Finsler billiard trajectories correspond to r -gons inscribed in M and having extremal Finsler length. The cyclic group \mathbb{Z}_r acts on these extremal polygons, and one counts the \mathbb{Z}_r -orbits. Using Morse and Lusternik–Schnirelmann theories, we prove that if $r \geq 3$ is prime, then the number of r -periodic Finsler billiard trajectories is not less than $(r-1)(d-2)+1$. We also give stronger lower bounds when M is in general position. The problem of estimating the number of periodic billiard trajectories from below goes back to Birkhoff. Our work extends to the Finsler setting the results previously obtained for Euclidean billiards by Babenko, Farber, Tabachnikov, and Karasev.

Key words: mathematical billiards; Finsler manifolds; magnetic billiards; Morse and Lusternik–Schnirelmann theories; unlabeled cyclic configuration spaces

2020 Mathematics Subject Classification: 37J45; 55R80; 70H12

*Dedicated to D. Fuchs on the occasion
of his 80th anniversary*

1 Introduction

The study of mathematical billiards goes back to G.D. Birkhoff who wrote in [11]:

... in this problem the formal side, usually so formidable in dynamics, almost completely disappears, and only the interesting qualitative questions need to be considered.

One of the main motivations for the study of billiards has been their relation with mathematical physics and statistical mechanics, namely, with the Boltzmann ergodic hypothesis. We refer to the books [15, 31, 40, 42] for various aspects of mathematical billiards.

1.1 Billiards in Euclidean geometry

A Birkhoff billiard table is bounded by a smooth strictly convex closed hypersurface M in \mathbb{R}^d . The billiard dynamical system describes the motion of a free particle inside M with elastic reflection off the boundary. That is, the point (billiard ball) moves with unit speed along a straight line until it hits the boundary M ; at the impact point, the normal component of the velocity instantaneously changes sign, while the tangential component remains the same, and the point continues its rectilinear motion. In dimension two, this is the familiar law of geometrical optics: the angle of incidence equals the angle of reflection.

An r -periodic billiard trajectory in M is an r -tuple of points (x_1, \dots, x_r) of M such that $x_i \neq x_{i+1}$ and the billiard reflection in M takes segment $x_{i-1}x_i$ to $x_i x_{i+1}$ for $i = 1, \dots, r$ (as usual, the indices are understood cyclically, that is, $x_{r+i} = x_i$).

The Dihedral group of symmetries of a regular r -gon D_r acts naturally on the set of all periodic billiard trajectories of period r by cyclically permuting the points and reversing the orientation. Thus, when counting periodic orbits, it is natural to count such dihedral orbits.

Problem 1.1. *Let $d \geq 2$ and $r \geq 2$ be integers, and let $M^{d-1} \subseteq \mathbb{R}^d$ be a smooth closed strictly convex hypersurface. Estimate below the number $N_E(M^{d-1}, r)$ of equivalence classes of periodic billiard trajectories of period r inside M^{d-1} modulo the action of the dihedral group D_r .*

The way we have formulated the problem, multiple trajectories are included into the count, so that for example a 2-periodic trajectory traversed thrice and a 3-periodic trajectory traversed twice, contribute to the number of 6-periodic trajectories. Of course, this issue is not a concern if r is a prime.

The first progress in addressing the question posed in Problem 1.1 was made by Birkhoff in 1927 [11]. He considered the case $d = 2$, and proved that there exist at least two r -periodic orbits with every rotation number coprime with r , which implies that $N_E(M^1, r) \geq 2\phi(\lfloor \frac{r}{2} \rfloor)$, where ϕ is the Euler totient function. We remark that this lower bound holds for the number of prime periodic trajectories. Birkhoff deduced his result from Poincaré's geometric theorem, that Poincaré published without proof shortly before his death and that Birkhoff proved a year later.

The concept of rotation number is not available in dimensions $d \geq 3$, but one can use the variational approach to the problem. Periodic billiard trajectories correspond to the critical points of the length function on r -tuple of points (x_1, \dots, x_r) , that is, on r -gons inscribed in M :

$$L(x_1, \dots, x_r) = \sum_{i=1}^r |x_i - x_{i+1}|. \tag{1.1}$$

This function is D_r -invariant.

The first to address Problem 1.1 for $d = 3$ was Babenko [6], whose approach was based on analyzing critical points of the length function L . Although his paper contained an error, the main idea was rescued and refined by Farber and Tabachnikov who established several lower bounds for $N_E(M^{d-1}, r)$.

When dealing with critical points of functions, one applies either the Morse theory or the Lusternik–Schnirelmann theory. The former usually gives stronger lower bounds on the number of critical points, but it applies to a more restrictive set of smooth functions, namely, Morse (or Morse–Bott) functions; the latter sometimes gives weaker lower bounds on the number of critical points, but it does not rely on genericity assumptions on the functions involved.

The variational methods subsequently produced a number of improved lower bounds in a sequence of papers by a number of authors:

- *Farber & Tabachnikov* [22, Theorem 1(B), p. 555] and [21, Theorem 3]. Let $d \geq 3$ be an integer, let $r \geq 3$ be an odd integer, and let M^{d-1} be a generic smooth closed strictly convex hypersurface. Then

$$N_E(M^{d-1}, r) \geq (r-1)(d-1).$$

This is a Morse-theoretical result, that is why M is assumed to be generic. The next results are Lusternik–Schnirelmann-theoretical, and they hold without the genericity assumption.

- *Farber & Tabachnikov* [21, Theorem 1(A), p. 554]. Let $d \geq 4$ be an integer, let $r \geq 3$ be an odd integer, and let M^{d-1} be a smooth closed strictly convex hypersurface. Then

$$N_E(M^{d-1}, r) \geq \lfloor \log_2(r-1) \rfloor + d - 1.$$

- *Farber* [20, Theorem 2, p. 589]. Let $r \geq 3$ be an odd prime, and let M^{d-1} be a smooth closed strictly convex hypersurface.

– If $d \geq 4$ is even, then

$$N_E(M^{d-1}, r) \geq r.$$

– If $d \geq 3$ is odd, then

$$N_E(M^{d-1}, r) \geq \frac{r+1}{2}.$$

- *Karasev* [29, Theorem 1, p. 424]. Let $d \geq 3$, let $r \geq 3$ be prime, and let M^{d-1} be a smooth closed strictly convex hypersurface. Then

$$N_E(M^{d-1}, r) \geq (r-1)(d-2) + 2.$$

Let us also mention Mazzucchelli’s paper [33] concerning the multiplicity of Birkhoff billiard periodic trajectories whose period is a power of a fixed prime number.

1.2 Billiards in Finsler geometry

Our goal is to extend the above described results to periodic trajectories in Finsler billiards.

From the point of view of physics, Finsler geometry describes the propagation of light in a medium that is not necessarily homogeneous or isotropic. The speed of light depends on the point and the direction, and is given by a smoothly varying norm on the tangent spaces to the medium thought of as a smooth manifold. We allow these norms to be asymmetric. These norms need not correspond to inner products, which is the case when the metric is Riemannian. To quote Chern’s description [14], “Finsler geometry is just Riemannian geometry without the quadratic restriction”.

The distance $f(A, B)$ between points A and B is defined as the least time it takes light to travel from A to B ; in general, $f(A, B) \neq f(B, A)$. The trajectories of light are Finsler geodesics. See Section 2.1 for precise definitions.

An example of a Finsler manifold is a Minkowski space, that is, a finite-dimensional normed space. Another example is a projective metric in a domain in the projective space, a Finsler metric whose geodesics are straight lines. Hilbert’s fourth problem asked to describe all projective Finsler metrics (a projective Riemannian metric is a metric of constant curvature), see, e.g., [13].

The first steps in the study of billiards in Finsler geometry were made by Gutkin and Tabachnikov in [27]. In this paper, the Finsler billiard reflection is defined (see Section 2.2 below) using the variational approach, and Minkowski billiards are studied in some detail. Minkowski billiards have close relations with convex geometry (Mahler's conjecture) and symplectic topology (symplectic capacities), see [4, 5].

We consider a domain in a Finsler manifold bounded by a smooth closed hypersurface M^{d-1} , a Finsler billiard table. We assume that M is *strictly convex* in the following sense:

- for every pair of points x, y inside M , there is a unique geodesic from x to y , and a unique geodesic from y to x , both contained inside M and distance-minimizing;
- M is quadratically convex: every geodesic tangent to M has second order contact with M , and not higher order.

In particular, these conditions imply that M is diffeomorphic to the sphere S^{d-1} . The convexity is an open condition. In the Minkowski setting, this condition is the same as in Euclidean space; see the end of Section 2.3 for the geometric meaning of convexity in the case of planar magnetic billiards.

We are interested in periodic Finsler billiard trajectories inside M . They still correspond to critical points of the analog of the length function (1.1)

$$\Lambda(x_1, \dots, x_r) = f(x_1, x_2) + f(x_2, x_3) + \dots + f(x_r, x_1),$$

however this function has less symmetry than in the Euclidean case: it is invariant under the cyclic permutations of the points, but not under the orientation reversal. Thus $\Lambda(x_1, \dots, x_r)$ is \mathbb{Z}_r -invariant, but not necessarily D_r -invariant. Accordingly, when counting periodic Finsler billiard trajectories, we count \mathbb{Z}_r -orbits.

Denote by $N_F(M^{d-1}, r)$ the number of equivalence classes of periodic Finsler billiard trajectories of period r inside M^{d-1} modulo the action of the cyclic group \mathbb{Z}_r .

1.3 Statement of main results

Our main result is as follows.

Main Theorem 1.2. *Let $d \geq 3$ be an integer and $r \geq 3$ be a prime. Consider Finsler billiard inside a smooth closed hypersurface M^{d-1} , satisfying the above formulated strict convexity assumptions. Then*

$$(A) \quad N_F(M^{d-1}, r) \geq (r-1)(d-2) + 1.$$

(B) *For a generic M ,*

- (1) *if d is even, then $N_F(M^{d-1}, r) \geq (r-1)d$;*
- (2) *if d is odd, then $N_F(M^{d-1}, r) \geq (r-1)(d-1)$.*

The general position assumption in case (B) is the assumption that the length $\Lambda(x_1, \dots, x_r)$ is a Morse function. The latter condition is generic in the sense that it holds in an open dense subset of strictly convex hypersurfaces, considered in the Whitney C^∞ topology. For Euclidean billiards, this is deduced in [22, Lemma 4.4] from an appropriate version of the multi-jet transversality theorem. A similar argument works in the Finsler case; we do not elaborate on it here.

Remark 1.3. The rate of growth of the numbers $N_F(M^{d-1}, r)$, provided by Theorem 1.2, is the same as in the above described results for Euclidean billiards: it is, roughly, rd . Since we

count \mathbb{Z}_r -orbits of periodic Finsler billiard trajectories, rather than D_r -orbits, one might expect the numbers $N_F(M^{d-1}, r)$ to be about twice as large as the numbers $N_E(M^{d-1}, r)$.

For example, consider Euclidean billiard inside a strictly convex closed smooth hypersurface, and switch on a weak magnetic field. One may expect each periodic trajectory in absence of the magnetic field to give rise to two periodic magnetic billiard trajectories, see Fig. 1.

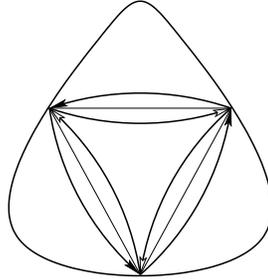


Figure 1. A 3-periodic billiard trajectory giving rise to two 3-periodic magnetic billiard trajectories.

In dimension two, this is indeed the case: the Finsler billiard map is an area preserving twist map, and it has two r -periodic trajectories for every rotation number k coprime with r . If the metric is symmetric, then the orbits corresponding to the rotation numbers k and $r - k$ differ only by the orientation, and they are counted as one, but in the asymmetric case, these are indeed different orbits.

We do not know how close the lower bounds of Theorem 1.2 are to being sharp. One may expect a notable difference between the reversible and non-reversible cases. In the related problem of closed geodesics, it is known that every Riemannian metric on the 2-sphere possesses infinitely many geometrically distinct closed geodesics [7, 25], but a Finsler metric may have precisely two distinct prime closed geodesics, as in the well known Katok example [30] (in which the two closed geodesics are the inverses of each other). Do similar examples exist for Finsler billiards?

2 From geometry of billiards to topology of cyclic configuration spaces

2.1 Introduction to Finsler geometry

We begin with a very brief introduction to Finsler geometry. For a thorough treatment, see [3, 8, 38].

A Finsler metric on a smooth manifold U is determined by a smooth non-negative fiberwise-convex Lagrangian function $L: TU \rightarrow [0, \infty)$, with the property that on each tangent space $T_y U$, L is positively homogeneous: $L(y, tv) = tL(y, v)$ for non-negative t and positive off the zero section. The restriction of L to any tangent space $T_y U$ gives the Finsler length of vectors in $T_y U$.

The vectors in $T_y U$ of unit Finsler length form a strictly convex hypersurface $I \subset T_y U$, called the *indicatrix*, which plays the role of the unit sphere in Riemannian geometry. We make the additional assumption that each indicatrix is quadratically convex. Specifying a smooth field of indicatrices on U is an equivalent method of defining a Finsler metric on U .

The Finsler metric on TU induces a notion of distance on the base manifold U . The length of a smooth curve $\gamma: [a, b] \rightarrow U$ is given by the integral

$$\text{Length}(\gamma) = \int_a^b L(\gamma(t), \gamma'(t)) dt.$$

The length of γ is independent of the parametrization, and a Finsler geodesic is an extremal of the length functional. In particular, a Finsler geodesic γ satisfies the Euler–Lagrange equations:

$$L_{vv}(\gamma(t), \gamma'(t)) \cdot \gamma''(t) + L_{vy}(\gamma(t), \gamma'(t)) \cdot \gamma'(t) = L_y(\gamma(t), \gamma'(t)).$$

In this formula we use a shorthand notation, so that L_v is the vector $(L_{v_1}, \dots, L_{v_n})$ and L_{vv} is the matrix $(L_{v_i v_j})$, etc.

For each pair of points x and y there is a corresponding *Finsler distance* $f(x, y)$, equal to the length of the shortest oriented geodesic from x to y . We stress that since geodesics are not necessarily reversible, this distance function f need not be symmetric, and therefore f is not a genuine metric on U .

The *figuratix* $J \subset T_y^*U$ is the “unit sphere of the cotangent space”, defined as follows: for a vector v in the indicatrix $I \subset T_yU$, there is a unique covector D_v , defined by the properties that $\text{Ker}(D_v) = T_vI$ and $D_v(v) = 1$. The map $D: I \rightarrow T_y^*U$ which maps v to D_v is the *Legendre transform*, and the image is the figuratrix J . The dual transform $D^*: J \rightarrow I$ is defined similarly. The Legendre transform is an involution: the composition of D and D^* is the identity map. Rightfully, the Legendre transform is a smooth bundle map from the indicatrix bundle to the figuratrix bundle, but we will often use the notation $D: I \rightarrow J$ when the basepoint y is understood.

2.2 The Finsler billiard map

Let U be a smooth d -dimensional Finsler manifold. Let $X \subset U$ be a compact d -dimensional submanifold with boundary $M = \partial X$. We assume that X and M satisfy the strict convexity assumptions formulated in Section 1.2. We will refer to X as a *billiard table*.

Let xy and yz be two oriented geodesic segments, where $x, z \in X$ are in the interior of the billiard table, and $y \in M$ is on the boundary. We say that yz is the *Finsler billiard reflection* of xy if y is a critical point of the distance function $f(x, \cdot) + f(\cdot, z): M \rightarrow [0, \infty)$. This relation is *not symmetric*, so it does not imply that yx is the billiard reflection of zy .

We describe the reflection law in the context of the Finsler setup following the treatment in [27]. Although the Finsler metric there is assumed to be symmetric, the reflection law is the same.

For each $y \in M$, let $p \in J \subset T_y^*U$ be the *conormal*, defined as the unit cotangent vector which vanishes on T_yM and is positive on the outward vectors. Let xy be an incoming geodesic and yz the reflected outgoing geodesic, corresponding respectively to tangent vectors u and v in I . The reflection xy to yz manifests as the following relation in the cotangent space.

Lemma 2.1 (Finsler billiard reflection law). *The covector $D_u - D_v$ is conormal to T_yM ; in particular $D_u - D_v = tp$ for some $t > 0$, see Fig. 2.*

Proof. The point $y \in M$ is a relative extremum of the function $f(x, \cdot) + f(\cdot, z)$, so the differential of this function is conormal to the tangent hyperplane T_yM . Let us compute this differential to show that it is $D_u - D_v$.

Fix a point $x \in X$, and consider the wave propagation from x . Let c be a non-singular point of the wave front \mathcal{F}_{t_0} such that the oriented geodesic segment xc of length t_0 is contained in the interior of X . The Finsler length function $f(x, \cdot)$ extends to a smooth function in a neighborhood of c . More precisely, for every point d sufficiently close to c , there exists a t near t_0 such that $d \in \mathcal{F}_t$. Let I and J represent the indicatrix and figuratrix at c , and let $u \in I$ be the Finsler unit vector along xc .

We claim that, at the point c , $df(x, \cdot) = D_u$. The wave front \mathcal{F}_{t_0} is a level set of $f(x, \cdot)$. Hence $df(x, \cdot)$ annihilates the tangent space to \mathcal{F}_{t_0} . By the Huygens principle, D_u is conormal

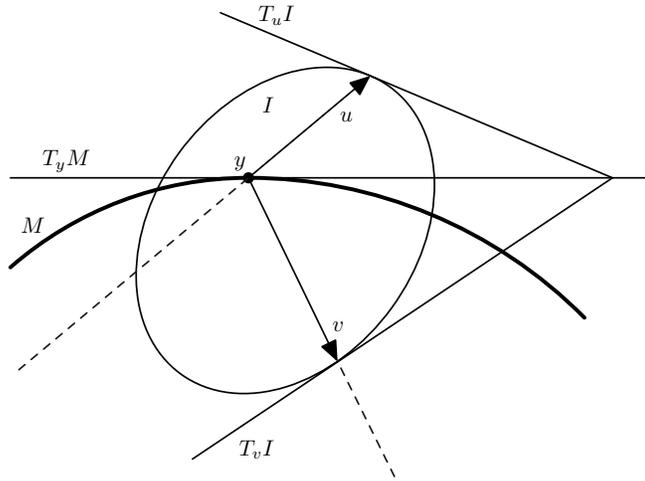


Figure 2. The Finsler billiard reflection in dimension two.

to this plane, hence proportional to $df(x, \cdot)$. It also follows from the Huygens principle that $df(x, \cdot)(u) = 1$, which is true of D_u by definition. Hence $df(x, \cdot) = D_u$.

Since the Finsler distance function is non-symmetric, we also need to understand $df(\cdot, z)$. Consider the oriented geodesic segment yz . Although zy is not necessarily a geodesic with respect to the Finsler structure L , it is a geodesic with respect to the “reverse” Finsler structure on M defined by the Lagrangian $\bar{L}(v) := L(-v)$. For the corresponding Finsler distance function \bar{f} we have the equality $\bar{f}(z, y) = f(y, z)$, and for the corresponding Legendre transform \bar{D} we have the equality $\bar{D}_{-v} = -D_v$ (by definition of the covector D_v). Therefore we have $df(\cdot, z) = d\bar{f}(z, \cdot) = \bar{D}_{-v} = -D_v$, where the middle equality follows from the paragraph above. This completes the proof. \blacksquare

2.3 Magnetic billiards as Finsler billiards

A popular billiard model that has been extensively studied in the last decades are magnetic billiards [9, 10, 26, 35, 41, 43, 44].

On a Riemannian surface, a magnetic field is given by a function $B(x)$, and the motion of a charged particle is described by the differential equation $\ddot{x} = B(x)J\dot{x}$, where J is the rotation of the tangent plane by $\pi/2$. This equation implies that the speed of the particle remains constant (the Lorentz force is perpendicular to the direction of motion). In the Euclidean plane, if the magnetic field is constant, the trajectories are circles of a fixed radius (Larmor circles).

In general, a magnetic field on a Riemannian manifold M is a closed differential 2-form β , and the magnetic flow is the Hamiltonian flow of the usual Hamiltonian $|p|^2/2$ on the cotangent bundle T^*M with respect to the twisted symplectic structure $\omega + \pi^*(\beta)$, where $\omega = dp \wedge dq$ is the standard symplectic form and $\pi: T^*M \rightarrow M$ is the projection.

Magnetic billiard describes the motion of a charged particle confined to a domain with elastically reflecting boundary. The reflection law is the same as for the usual billiards, with zero magnetic field: the angle of incidence equals the angle of reflection. Following [41], one can interpret magnetic billiards as Finsler ones.

Let us assume that the magnetic 2-form is exact: $\beta = d\alpha$ for some differential 1-form on a Riemannian manifold M^n . Then the magnetic flow admits a Lagrangian formulation with the Lagrangian function

$$\bar{L}(x, v) = \frac{1}{2}|v|^2 + \alpha(x)(v),$$

where $x \in M$, $v \in T_x M$, and $|v|$ is the Riemannian norm of the tangent vector v .

We want to consider the motion with unit speed, that is, to fix an energy level. Following the Maupertuis principle, replace the Lagrangian \bar{L} with

$$L(x, v) = |v| + \alpha(x)(v). \quad (2.1)$$

Assume that the magnetic field is weak enough so that $L(x, v) > 0$ for all non-zero tangent vectors v and all points x , that is, we assume that $|\alpha(x)| < 1$ everywhere. Then formula (2.1) defines a non-symmetric Finsler metric.

Lemma 2.2 (magnetic billiard reflection law). *The indicatrix of the magnetic Finsler metric (2.1) is an ellipsoid of revolution with a focus at the origin, and the Finsler billiard reflection law coincides with the usual one: the angle of incidence equals the angle of reflection.*

Proof. Fix a point x and consider the tangent space at this point. The tangent space is Euclidean, and the indicatrix I is given by the equation $|v| + \alpha(v) = 1$. Let e be the tangent vector dual to the covector α , that is, $\alpha(v) = e \cdot v$. We can choose an orthonormal basis so that $e = (t, 0, \dots, 0)$ with $t < 1$.

The equation of the indicatrix is $L(x, v) = 1$, or $x_1^2 + \dots + x_n^2 = (1 - tx_1)^2$, that is,

$$(1 - t^2)^2 \left(x_1 + \frac{t}{1 - t^2} \right)^2 + (1 - t^2)(x_2^2 + \dots + x_n^2) = 1.$$

This is the equation of an ellipsoid of revolution obtain by revolving the ellipse

$$\frac{(x_1 + c)^2}{a^2} + \frac{x_2^2}{b^2} = 1,$$

where

$$a^2 = \frac{1}{(1 - t^2)^2}, \quad b^2 = \frac{1}{(1 - t^2)}, \quad c = \frac{t}{1 - t^2},$$

and hence $c^2 = a^2 - b^2$. Therefore this ellipse has a focus at the origin.

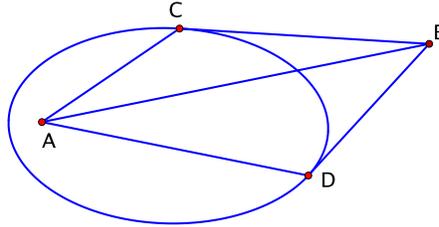


Figure 3. Point A is a focus of an ellipse; then $\angle BAC = \angle BAD$.

The second statement of the lemma reduces to a known geometric property of conics depicted in Fig. 3, see, e.g., [2]. Compare with the Finsler reflection law, Fig. 2. ■

We assume that the convexity conditions of Section 1.2 hold for magnetic billiards. This implies that the magnetic field is weak enough. For example, if the billiard table is a planar domain bounded by a smooth strictly convex curve and the magnetic field is constant, then the Larmor radius is greater than the greatest radius of curvature of the boundary curve. This is one of the three regimes in magnetic billiards described in [35], in which “motion is qualitatively similar to the field-free case”, albeit not time-reversible.

2.4 Morse-theoretic approach to Finsler billiards

We now discuss the necessary Finsler geometry and topology to apply results from Morse and Lusternik–Schnirelmann theories to the problem of periodic Finsler billiard trajectories. Similar preparation work is easier to do in the Euclidean case; it was the content of [22, Section 4].

The *ordered cyclic configuration space* of r consecutively distinct points on M is the space

$$G(M, r) := \{(x_1, \dots, x_r) \in M^{\times r} : x_i \neq x_{i+1} \text{ for all } i\},$$

where by convention $x_{r+1} = x_1$. As mentioned in the Introduction, we consider the *length function*

$$\Lambda: G(M, r) \longrightarrow \mathbb{R}, \quad \Lambda(x_1, \dots, x_r) = f(x_1, x_2) + f(x_2, x_3) + \dots + f(x_r, x_1).$$

The function Λ is smooth and \mathbb{Z}_r -equivariant, but in contrast with Euclidean billiards, Λ need not be D_r -equivariant. By definition of the Finsler billiard reflection for geodesic rays (Section 2.2), the r -periodic Finsler billiard orbits are precisely the critical points of Λ .

We would like to use this fact by applying Morse and Lusternik–Schnirelmann theories to obtain a lower bound for the number of periodic Finsler billiard orbits; however, these theories cannot be applied directly because $G(M, r)$ is not a closed manifold. We will show that for sufficiently small ε we can replace $G(M, r)$, without affecting the topology, by the following compact manifold with boundary:

$$G_\varepsilon(M, r) := \left\{ (x_1, \dots, x_r) \in M^{\times r} \mid \prod_{i=1}^r f(x_i, x_{i+1}) \geq \varepsilon > 0 \right\},$$

where again the indices are understood cyclically. Similar to the Euclidean case [22, Proposition 4.1], we establish the following proposition in the Finsler setup.

Proposition 2.3. *If $\varepsilon > 0$ is sufficiently small, then*

- (1) $G_\varepsilon(M, r)$ is a smooth manifold with boundary;
- (2) the inclusion $G_\varepsilon(M, r) \subset G(M, r)$ is a \mathbb{Z}_r -equivariant homotopy equivalence;
- (3) all critical points of $\Lambda: G(M, r) \longrightarrow \mathbb{R}$ are contained in the interior of $G_\varepsilon(M, r)$;
- (4) at every critical point of $\Lambda|_{\partial G_\varepsilon(M, r)}$, the differential $d\Lambda$ is positive on inward vectors.

This proposition makes it possible to apply \mathbb{Z}_r -equivariant Morse and Lusternik–Schnirelmann theories to the length function Λ on the manifold with boundary $G_\varepsilon(M, r)$. We shall restrict ourselves to the case when r is prime. Then the group \mathbb{Z}_r acts freely on $G_\varepsilon(M, r)$, and \mathbb{Z}_r -equivariant Morse theory reduces to Morse theory on the quotient manifold $G_\varepsilon(M, r)/\mathbb{Z}_r$.

Due to item (4) of Proposition 2.3, the topological lower bounds on the number of critical points of Λ , such as the sum of Betti numbers or the Lusternik–Schnirelmann category, come from the topology of $G_\varepsilon(M, r)/\mathbb{Z}_r$. This space has the same topology as the cyclic configuration space $G(S^{d-1}, r)/\mathbb{Z}_r$, due to item (2), and the fact that M is topologically the sphere. We refer to [32] for the Morse theory on manifolds with boundary.

2.5 Technical bounds for Finsler billiards

We start with a number of technical lemmas working toward the proof of Proposition 2.3.

The boundary $\partial G_\varepsilon(M, r)$ is the level set (at ε^2) of the smooth function

$$F: M^{\times r} \longrightarrow \mathbb{R}, \quad F(x_1, \dots, x_r) = \prod_{i=1}^r f(x_i, x_{i+1})^2,$$

and $F^{-1}(0) = M^{\times r} \setminus G(M, r)$ is a critical level. The first two items are therefore a consequence of the following lemma.

Lemma 2.4. *There exists a constant $\delta > 0$ such that the interval $(0, \delta)$ consists of regular values of F .*

We offer some geometric intuition for this statement. As indicated above, the critical points of the length function Λ are precisely the periodic Finsler billiard orbits. Similarly, we may think of the critical points of the function F as the periodic orbits of some “unusual” billiard trajectory, which we will call the F -billiard trajectory, and for which the reflection law is given by Lemma 2.7 below. In this terminology, Lemma 2.4 claims the existence of δ such that any r -periodic F -billiard orbit (x_1, \dots, x_r) satisfies $F(x_1, \dots, x_r) \geq \delta$.

Assume that no such δ exists. Then we can find a closed F -billiard orbit such that one of the edge lengths $\ell_i := f(x_{i-1}, x_i)$ is arbitrarily small. The contradiction will arise in Section 2.7, where we show the following two statements:

- (1) If one edge of a closed F -orbit is “arbitrarily short”, then *all* of the edges are “arbitrarily short”.
- (2) A closed F -orbit cannot have all edges “arbitrarily short”.

We will not explicitly define *arbitrarily short*, but a suitable quantity could be determined in terms of the period r and certain “curvature” quantities, which depend on the Finsler geometry of M and on the geometry of the indicatrix bundle. A similar shortness statement was proven in the Euclidean case [22]; however, those curvature estimates rely on symmetry of the inner product, symmetry of the unit tangent spheres, and also some trigonometry; none of which we have at our disposal in the Finsler setting. We treat these subtleties in the following Sections 2.5.1–2.5.3.

2.5.1 A “curvature” bound for Finsler manifolds

In the Finsler setting, we would like to formalize the intuitive idea that for $y \in M$, “a geodesic segment yz is short if and only if the corresponding tangent vector at y is almost tangent to M ”.

This is easy to establish in the Euclidean case [22] as follows. The boundary M of the billiard table is a strictly convex hypersurface in Euclidean space \mathbb{R}^d . By strict convexity of M , there exist positive numbers $\rho < R$ such that for every $y \in M$, there exist two spheres tangent to M at y , of radii ρ and R , such that the sphere of radius R contains M , and the sphere of radius ρ is contained in M . Let n be the outward unit normal at y and let $v \in T_y \mathbb{R}^d$ be a unit tangent vector with $\langle v, n \rangle < 0$, and follow the geodesic from y in the direction of v until colliding again with the boundary, say at $z \in M$. Then the numbers ρ and R satisfy

$$\rho < \frac{|y - z|}{-2\langle v, n \rangle} < R.$$

In particular, *the measurements $g(v) := |y - z|$ and $-p(v) = -\langle v, n \rangle$ are of the same order*. We will use the notation $g \sim p$ for such a statement.

Now let U be a Finsler manifold and M a smooth, closed hypersurface, quadratically convex with respect to the Finsler geodesics. For $y \in M$, let $p \in J$ represent the unit covector which is conormal to M at y and positive on outward vectors. Given $v \in I \subset T_y U$ such that $p(v) < 0$, let $z \in M$ be the first collision with M of the geodesic ray emanating from y in the direction v . Define $g(v) = f(y, z)$.

Similarly, given $u \in I \subset T_y U$ such that $p(u) > 0$, let $x \in M$ be the point such that the oriented geodesic ray xy enters y with direction u . Define $g(u) = f(x, y)$.

Remark 2.5. To verify that such a point x exists, note that the oriented ray xy is a geodesic if and only if the oriented ray yx is a geodesic with respect to the “reverse” Finsler structure defined by the Lagrangian $\bar{L}(u) := L(-u)$. The point x can be defined as the first collision with M of the reverse geodesic ray emanating from y in the direction $-u$.

Lemma 2.6. *There exist positive constants $\rho < R$ such that, for all $y \in M$, all $v \in I \subset T_y U$ with $p(v) < 0$, and all $u \in I \subset T_y U$ with $p(u) > 0$, we have*

$$\rho < \frac{f(y, z)}{-p(v)} < R \quad \text{and} \quad \rho < \frac{f(x, y)}{p(u)} < R.$$

Equivalently, the functions g and p have the same order, $g \sim p$.

The following proof is due to Sergei Ivanov.

Proof. We will show the left inequalities, corresponding to vectors v emanating from y . The right inequalities follow similarly.

By compactness of M , it suffices to show existence for a single point $y \in M$. In fact, it is enough to show that the lemma holds near y : that is, for $v \in I \subset T_y U$ with $p(v)$ near 0. This is due to the following fact: for all v such that $p(v)$ is sufficiently away from 0, the quantity $\frac{f(y, z)}{-p(v)}$ is bounded away from zero. Indeed, as $f(y, z)$ tends to zero, that is, point z tends to point y , the vector along the geodesic yz tends to a tangent vector to M at y , that is, $p(v)$ also tends to zero.

Consider smooth coordinates (y_1, \dots, y_d) near y , such that a neighborhood of y in M is given by the equation $y_d = 0$. Let $\gamma(t) = (\gamma_1(t), \dots, \gamma_d(t))$ be the unit-speed geodesic passing through y at time 0 and with unit tangent vector $v \in I$ (see Fig. 4, left).

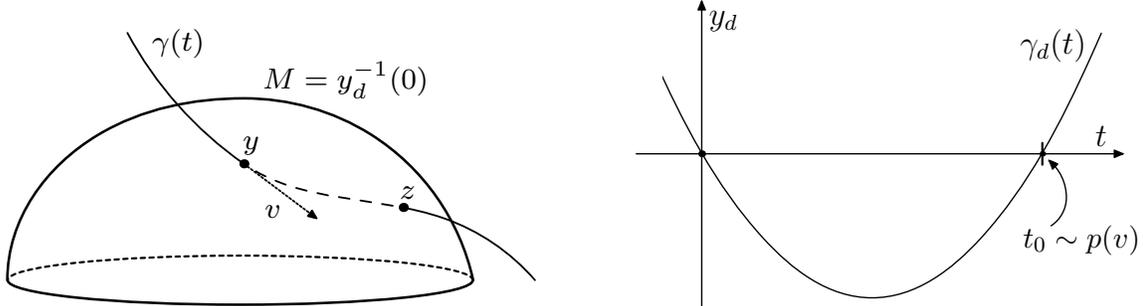


Figure 4. (Left): The unit speed geodesic $\gamma(t) = (\gamma_1(t), \dots, \gamma_d(t))$ passes through $y \in M$ at time 0 with tangent vector v and again at $z \in M$ at time $t_0 \sim p(v)$. (Right): The d^{th} component of γ , shown to the second order.

In case $v \in I \cap T_y M$, we have $\gamma_d(0) = 0$ and $\gamma'_d(0) = 0$, so by quadratic convexity, $\gamma''_d(0)$ is bounded away from 0 for all $v \in I \cap T_y M$. Therefore, for v almost tangent to M (i.e., with $p(v) < 0$ sufficiently near 0), we have $\gamma''_d(0) \sim 1$. In addition, for v almost tangent to M , we have $p(v) \sim \gamma'_d(0) < 0$.

Now writing $\gamma_d(t)$ up to second order yields

$$\gamma_d(t) = \gamma_d(0) + \gamma'_d(0)t + \frac{1}{2}\gamma''_d(0)t^2,$$

as depicted in Fig. 4, right. Thus $\gamma_d(t)$ will meet zero again at a time

$$t_0 = -2\frac{\gamma'_d(0)}{\gamma''_d(0)} \sim p(v).$$

Since t_0 is the length $f(y, z) = g(v)$, we conclude that $g \sim p$, completing the proof. ■

2.5.2 Reflection law for the F -billiard

Just as the critical points of the length function Λ correspond to periodic orbits of the Finsler billiard, we think of the critical points of F as the periodic orbits of some “unusual” billiard trajectory determined by the function F . In fact, what we are really interested in are the critical points of F ; we use this billiard terminology because it is intuitive and suggestive.

So suppose that $x, y, z \in M$ are points such that the oriented geodesic segment yz is the F -reflection of the oriented geodesic segment xy . Let $u, v \in I \subset T_y U$ be the tangent vectors which correspond, respectively, to xy and yz . Let $p \in J \subset T_y^* U$ represent the outward-pointing unit conormal, and let $n = D_p^*$ represent the outward pointing unit normal. Then the F -billiard has the following reflection law.

Lemma 2.7. *The F -billiard reflection law is given by the following cotangent relation:*

$$\frac{D_u}{f(x, y)} - \frac{D_v}{f(y, z)} = tp, \quad t > 0.$$

Proof. The point $y \in M$ is a relative extremum of the function $F(\cdot) := f(x, \cdot)^2 f(\cdot, z)^2$, so the differential of this function at y is conormal to the tangent hyperplane $T_y M$. Therefore the differential of the function

$$\frac{1}{2} \ln F(\cdot) = \ln f(x, \cdot) + \ln f(\cdot, z)$$

at y is also conormal to the tangent hyperplane $T_y M$. We compute

$$\frac{1}{2} d(\ln F(\cdot)) = d(\ln f(x, \cdot) + \ln f(\cdot, z)) = \frac{df(x, \cdot)}{f(x, \cdot)} + \frac{df(\cdot, z)}{f(\cdot, z)}.$$

Recall from the proof of Lemma 2.1 that at the point y , $df(x, \cdot) = D_u$ and $df(\cdot, z) = -D_v$. Therefore, at the point y , we have

$$\frac{D_u}{f(x, y)} - \frac{D_v}{f(y, z)} = tp,$$

as desired. ■

We will also make use of the following consequence.

Lemma 2.8. *If $v \in I \subset T_y M$ is the F -reflection of $u \in I \subset T_y M$, then either the linear subspaces $T_u I, T_v I, T_y M \subset T_y U$ are equal, or they intersect in a subspace of codimension 2.*

Proof. First of all, we clarify the statement. The indicatrix I is a hypersurface in the vector space $T_y U$, hence its tangent spaces can be considered as vector subspaces of $T_y U$. It is in this sense that we intersect them, and $T_y M$, in the ambient space $T_y U$. This remark applies to other similar arguments elsewhere in the paper.

Now, by the reflection law, the covectors D_u, D_v , and p are linearly dependent, hence span a subspace of at most two dimensions. Therefore, the intersection of their kernels, $T_u I \cap T_v I \cap T_y M$, has codimension at most two. ■

Given $u \in I \subset T_y U$ with $p(u) > 0$, let v be the F -reflection of u . Suppose that $u \neq n$, so that the hyperplanes $T_u I, T_v I$, and $T_y M$ are not parallel. Then, from Lemma 2.8, the intersection of the linear hyperplanes $T_u I \cap T_v I$ is a codimension 1 subspace of $T_y M$, hence there is a unique vector $w = w(u) \in T_y M \cap I$ for which the kernel of D_w contains $T_u I \cap T_v I$ and such that $D_u(w) > 0$. Now applying the (cotangent) relation of Lemma 2.7 to the vector w yields:

$$\frac{D_u(w)}{f(x, y)} = \frac{D_v(w)}{f(y, z)}. \tag{2.2}$$

Note that $D_v(w) > 0$, since the denominators are positive and $D_u(w) > 0$ by assumption.

2.5.3 Global indicatrix bound

Since the Finsler norm need not arise from an inner product, it is useful to develop some method of comparing two vectors $u, v \in I$. One non-symmetric idea is to apply the Legendre transform to obtain a covector D_u which then acts on v . We have a general lower bound by compactness and a specific upper bound by strict convexity of the indicatrices.

Lemma 2.9. *There exists a positive constant $K \geq 1$ such that, for every triple $y \in M$, $u, v \in I \subset T_y U$, one has $-K \leq D_u(v) \leq 1$. In particular, $D_u(v) = 1$ if and only if $u = v$.*

We will also require a more refined bound. For $y \in M$, consider $u \in I \subset T_y U$ such that $p(u) > 0$ and $u \neq n$. Let $v \in I \subset T_y U$ be the F -reflection of u and let w be as described in Section 2.5.2. Although $w = w(u)$ is not defined when $u = n$, any sequence $u_i \rightarrow n$ has $D_{u_i}(w_i) \rightarrow 0$ and $D_{v_i}(w_i) \rightarrow 0$, so that the maps $u \mapsto D_u(w)$ and $u \mapsto D_v(w)$ extend continuously to $u = n$.

Lemma 2.10. *There exists a positive constant $k < 1$ such that for all $y \in M$ and all $u \in I \subset T_y U$ with $p(u) > 0$,*

$$k < D_u(w) + p(u) \quad \text{and} \quad k < D_v(w) - p(v).$$

Proof. The numbers $D_u(w)$, $D_v(w)$, $p(u)$, and $-p(v)$ are all positive. The first two are bounded away from 0 except when u is almost parallel to n , and the last two are bounded away from 0 except when u is almost tangent to M . ■

2.6 The F -billiard reflection preserves shortness and almost-tangency

With all of the technical bounds obtained in Section 2.4, we are ready to study the F -billiard reflection in more detail. We show that short geodesics F -reflect to short geodesics, and that almost-tangent vectors F -reflect to almost-tangent vectors.

In particular, suppose that $u \in I \subset T_y U$, let $v \in I \subset T_y U$ be the F -reflection of u , and let x and z be points in M such that u and v correspond, respectively, to the oriented geodesic segments xy and yz . In this sense, we may consider v , x , and z as functions of u . Let $p \in J \subset T_y^* U$ be the outward pointing unit conormal and let $n = D_p^*$ be the outward pointing unit normal. Our goal is to show the following:

$$p(u) \sim f(x, y) \sim f(y, z) \sim p(v).$$

Note that the outer equivalences were already established in Lemma 2.6.

We make use of the positive constants ρ and R defined by Lemma 2.6, as well as the positive constants K and k determined in Lemmas 2.9 and 2.10.

Lemma 2.11. *If $v \in I \subset T_y U$ is the F -reflection of $u \in I \subset T_y U$, with $p(u) > 0$, then*

$$\frac{k\rho}{K(\rho + R)} \leq \frac{f(x, y)}{f(y, z)} \leq \frac{K(\rho + R)}{k\rho}.$$

In particular, the measurements $f(x, y)$ and $f(y, z)$ are equivalent: $f(x, y) \sim f(y, z)$.

Proof. We prove by contradiction. Assume that the left inequality fails, so that

$$\frac{f(x, y)}{f(y, z)} < \frac{k\rho}{K(\rho + R)}. \tag{2.3}$$

Then the F -reflection law applied to w (2.2), combined with (2.3) and then Lemma 2.9 twice, yields

$$D_u(w) = \frac{f(x, y)}{f(y, z)} D_v(w) < \frac{k\rho}{K(\rho + R)} D_v(w) \leq \frac{k\rho}{K(\rho + R)} \leq \frac{k\rho}{\rho + R}. \quad (2.4)$$

On the other hand, Lemma 2.6, combined with (2.3) and then Lemma 2.9, yields

$$p(u) < -\frac{kR}{K(\rho + R)} p(v) < \frac{kR}{\rho + R}. \quad (2.5)$$

But the sum of (2.4) and (2.5) gives a contradiction; the right side is equal to k , while the left side is greater than k by Lemma 2.10.

The right inequality of Lemma 2.11 can be shown similarly. \blacksquare

Lemma 2.12. *The F -billiard reflection $u \mapsto v$ can be extended continuously to $I \cap T_y M$. Moreover, this extension is the identity map on $I \cap T_y M$.*

Proof. This lemma clearly holds for the ordinary Finsler billiard reflection law in Lemma 2.1. In the F -billiard case, by Lemmas 2.6 and 2.11, we have $p(u) \sim f(x, y) \sim f(y, z) \sim p(v)$, so if a sequence $u_n \subset I$ tends to $T_y M$, so does the sequence of F -reflections v_n .

From Lemma 2.8, we have $\text{Ker } D_{u_n} \cap T_y M = \text{Ker } D_{v_n} \cap T_y M$ for every n . Therefore, if $u_n \rightarrow u \in I \cap T_y M$, then $\lim (\text{Ker } D_{u_n}) \cap T_y M$ is some codimension 1 subspace $W \subset T_y M$, and we must have $W = \text{Ker } D_v \cap T_y M$ for any continuously-defined F -reflection v of u . Since we also must have $v \in I \cap T_y M$, there are only two candidates for v : u itself and the ‘‘opposite’’ vector u' , for which $D_u(u') < 0$. Now, by definition of the unit vector $w_n \in T_y M$ (see Section 2.5.2), $u = \lim w_n$. We have $D_{v_n}(w_n) > 0$ for all n , so $D_v(u)$ must be nonnegative. Hence $v = u$. \blacksquare

We are now ready to prove Lemma 2.4.

2.7 Nonexistence of short periodic F -billiard trajectories

Proof of Lemma 2.4. Assume that (x_1, \dots, x_r) is an r -periodic F -billiard orbit. Let $u_i, v_i \in I \subset T_{x_i} U$ represent the tangent vectors which correspond, respectively, to the oriented geodesic segments $x_{i-1}x_i$ and $x_i x_{i+1}$. Let $p_i \in J \subset T_{x_i}^* U$ be the outward pointing unit conormal and let $n_i = D_{p_i}^*$ be the outward pointing unit normal. Let $\ell_i = f(x_{i-1}, x_i)$. We aim to show that no ℓ_i can be arbitrarily small. From Lemma 2.11 we obtain the edge comparison for any i and j :

$$\frac{\ell_i}{\ell_j} < \left(\frac{K(\rho + R)}{k\rho} \right)^r.$$

In particular, this establishes that $\ell_i \sim \ell_j$, so that if ℓ_i is small for any i , every ℓ_j is also small.

Now, seeking a contradiction, we assume that a critical r -gon (x_1, \dots, x_r) has all edges arbitrarily short. Then each u_i is almost tangent, so by Lemma 2.12, $u_i \sim v_i$ in any auxiliary metric on M . In particular, for any auxiliary Euclidean metric on M , (x_1, \dots, x_r) is an r -gon with all exterior angles small. The contradiction will arise once we show the following statement which is a discrete analog of the theorem that the total curvature of a spacial closed curve is not less than 2π (see, e.g., [23]):

The sum of the exterior angles of an r -gon in Euclidean space \mathbb{R}^q is at least 2π .

Let a_1, \dots, a_r be oriented vectors representing the edges of the polygon. The Gauss map sends these vectors to some points A_1, \dots, A_r on S^{q-1} , and the great circle segment connecting A_i and A_{i+1} has length equal to the exterior angle at the vertex connecting a_i and a_{i+1} (here the indices are cyclic). Thus it is enough to show that the perimeter of this spherical polygon

(A_1, \dots, A_r) is at least 2π . By the Crofton formula, for this it is enough to show that the polygon intersects every great sphere S^{q-2} .

Choose a great sphere $S = S^{q-2} \subset S^{q-1}$ and let $P = \mathbb{R}^{q-1} \subset \mathbb{R}^q$ be the corresponding hyperplane. If some vector a_i is contained in P , then S contains A_i . Otherwise, translate $P \subset \mathbb{R}^q$ so that it sits as a supporting hyperplane to the polygon (a_1, \dots, a_r) , say at the vertex connecting edge a_i to edge a_{i+1} . In this case the points A_i and A_{i+1} are separated by the great sphere S , so the great circle segment connecting A_i to A_{i+1} must intersect S . ■

2.8 Critical points of the restricted length function

With the proof of Lemma 2.4 complete, we turn our attention to the fourth item of Proposition 2.3, that if some r -gon $x = (x_1, \dots, x_r) \in \partial G_\varepsilon(M, r)$ is a critical point of $\Lambda|_{\partial G_\varepsilon(M, r)}$, then $d\Lambda_x$ is positive on inward vectors. Since $\partial G_\varepsilon(M, r)$ is a level set of F , the differential dF vanishes precisely on vectors tangent to $\partial G_\varepsilon(M, r)$. Then, since $dF_x = t d\Lambda_x$ on $T_x G_\varepsilon(M, r)$, the sign of $d\Lambda_x$ is the same for all inward pointing tangent vectors at point x . Therefore, it is enough to show that $d\Lambda_x$ is positive on a single inward vector $V \in T_x G_\varepsilon(M, r)$.

For the remainder of this section we let $x = (x_1, \dots, x_r)$ be a critical point as discussed above, and we let $u_i, v_i \in T_{x_i} U$ be unit vectors tangent to the geodesic rays $x_{i-1}x_i$ and $x_i x_{i+1}$. We will drop the subscript on the differentials $d\Lambda$ and dF , since we will only be discussing them at the point x . We will continue to use the notation n_i for the outward-pointing unit normal vector at x_i and $p_i = D_{n_i}$. We will use ℓ_i to represent the Finsler distance $f(x_{i-1}, x_i)$.

We recall the differentials:

$$d\Lambda = (D_{u_1} - D_{v_1}, \dots, D_{u_r} - D_{v_r}),$$

and

$$\frac{1}{2}d(\ln F) = \left(\frac{D_{u_1}}{\ell_1} - \frac{D_{v_1}}{\ell_2}, \dots, \frac{D_{u_r}}{\ell_r} - \frac{D_{v_r}}{\ell_1} \right).$$

To show that there exists $V \in T_x G_\varepsilon(M, r)$ for which both differentials are positive, it is enough to find a vector $\nu \in T_{x_i} M$, for some i , such that $D_{u_i}(\nu) > 0$ and $D_{v_i}(\nu) < 0$. Indeed, we can then let $V \in T_x G_\varepsilon(M, r)$ be the vector whose only nonzero component is ν .

The structure of this section is to assume that no such ν exists and analyze the consequences. The following lemma is the first step of this process.

Lemma 2.13. *Suppose that for some fixed i , there is no $\nu \in T_{x_i} M$ satisfying $D_{u_i}(\nu) > 0$ and $D_{v_i}(\nu) < 0$. Then there exist a_i and b_i , both positive or both negative, such that*

$$a_i D_{u_i} - b_i D_{v_i} = p_i. \tag{2.6}$$

Proof. The hypothesis implies that $\ker(D_{u_i}) \cap T_{x_i} M = \ker(D_{v_i}) \cap T_{x_i} M$, hence the covectors D_{u_i} and D_{v_i} are positive-proportional when restricted to $T_{x_i} M$. That is, there exist a_i and b_i , both positive or both negative, such that either

$$a_i D_{u_i} - b_i D_{v_i} = 0 \quad \text{or} \quad a_i D_{u_i} - b_i D_{v_i} = p_i.$$

The former is impossible since it implies that $u_i = v_i$. ■

It follows from Lemma 2.13 that the three covectors D_{u_i} , D_{v_i} , and p_i span a 2-plane P , hence there exists a unique vector $w_i \in T_{x_i} M \cap I$ such that $D_{w_i} \in P$ and $D_{u_i}(w_i) > 0$. In the next lemma we develop a universal constant, which allows us to compare D_{u_i} with $p_i(u_i)$ and D_{v_i} with $p_i(v_i)$, for any u_i and v_i satisfying a reflection law as in (2.6). We observe the similarities in (2.6) to both the Finsler reflection law and the F -reflection law, and we note that w_i is an analogue of the vector w introduced after Lemma 2.8 for the F -billiard reflection.

Lemma 2.14. *There exist constants $C > 0$ and $m \in \mathbb{N}$, such that, for all $x \in M$, and for all $w \in T_x M \cap I$, if $P = \text{Span}\{p, D_w\}$, and if $u \in I$ is any vector with $D_u \in P$ and $D_u(w) > 0$, then*

$$\frac{1}{C}p(u)^{2m} \leq 1 - D_u(w) \leq Cp(u)^{2m}.$$

Proof. By compactness of M and $T_x M \cap I$ it is enough to show the existence of C at a single point x and for a single 2-plane P . Let $\Gamma \subset I$ be the open curve consisting of those points u described in the statement of the lemma. We claim that $p|_\Gamma: \Gamma \rightarrow \mathbb{R}$ is injective. Indeed, suppose that $p(u) = p(v)$ for $u, v \in \Gamma$. We may write $aD_u - bD_v = p$; here a and b have the same sign, since $D_u(w)$ and $D_v(w)$ are both positive and $p(w) = 0$. Applying this relation to u and v yields

$$a - bD_v(u) = p(u) = p(v) = aD_u(v) - b,$$

therefore $a(1 - D_u(v)) = -b(1 - D_v(u))$, contradicting that a and b have the same sign.

Now consider the map $h: p(\Gamma) \rightarrow \mathbb{R}$ given by $p(u) \mapsto D_u(w)$. Then $h(0) = D_w(w) = 1$, and $h'(0) = 0$ since $D_u(w)$ is maximized at $u = w$. Therefore the first nonzero derivative of h has even order and is negative; in particular, $1 - D_u(w) \sim p(u)^{2m}$ for some m . ■

It follows from quadratic convexity of the indicatrix that $m = 1$, but we omit the details here. We are most interested in the following form of Lemma 2.14.

Corollary 2.15. *For any $x \in M$ and any $u, v \in I \subset T_x U$ such that $p(u) > 0$, $p(v) < 0$, and $aD_u - bD_v = p$, where a and b have the same sign, the following inequality holds:*

$$\frac{p(v)^{2m}}{C^2 p(u)^{2m}} \leq \frac{1 - D_v(w)}{1 - D_u(w)} \leq C^2 \frac{p(v)^{2m}}{p(u)^{2m}},$$

where $w \in T_x M$ is the unique unit vector with $D_w \in \text{Span}\{D_u, D_v\}$ and $D_u(w) > 0$.

We will show the existence of an appropriate vector ν in two separate cases: the first, when all side lengths of the r -gon are sufficiently short; and the second, when there is at least one long side. The following lemma treats the first case.

Lemma 2.16. *There exists $\eta > 0$ such that, if $p_i(u_i) < \eta$ and $-p_i(v_i) < \eta$ for all $i \in \{1, \dots, r\}$, then there exists an index i and a vector $\nu \in T_{x_i} M$, such that $D_{u_i}(\nu) > 0$ and $D_{v_i}(\nu) < 0$.*

Proof. Assume that no such ν exists, then equation (2.6), with a_i and b_i both positive or both negative, holds at every vertex i . We will use the ‘‘all small’’ hypothesis of the lemma to arrive at a contradiction. We first focus on a single vertex and drop the subscript i . Apply (2.6) to vectors u and v to obtain

$$a - bD_v(u) = p(u), \tag{2.7}$$

$$aD_u(v) - b = p(v). \tag{2.8}$$

Suppose that $\eta = \frac{\delta}{2+2K}$, where $\delta > 0$ is a small unspecified number and $K \geq 1$ is the constant determined in Lemma 2.9. We claim that $D_u(v)$ and $D_v(u)$ cannot be negative.

First, if a and b are negative, then $D_u(v)$ and $D_v(u)$ are both positive because $p(u) > 0$, $p(v) < 0$. Otherwise, if a and b are positive, and also if $D_v(u)$ is negative, then

$$a = p(u) + bD_v(u) < p(u),$$

$$b = -p(v) + aD_u(v) < -p(v) + p(u)D_u(v) < -p(v) + p(u),$$

and so a and b are both less than 2η . The same bounds hold if $D_u(v)$ is negative.

Now apply (2.6) to the vector n to obtain

$$1 = aD_u(n) - bD_v(n) < a - bD_v(n) < 2\eta + bK < 2\eta(1 + K) = \delta,$$

contradicting the assumption that $D_v(u)$ or $D_u(v)$ is negative. Now, solve for a and b using equations (2.7) and (2.8) to obtain

$$\begin{aligned} a &= \frac{p(u) - p(v)D_v(u)}{1 - D_u(v)D_v(u)} < \frac{2\eta}{1 - D_u(v)D_v(u)}, \\ b &= \frac{-p(v) + p(u)D_u(v)}{1 - D_u(v)D_v(u)} < \frac{2\eta}{1 - D_u(v)D_v(u)}. \end{aligned}$$

Again we apply (2.6) to the vector n to obtain

$$1 = aD_u(n) - bD_v(n) < \frac{2\eta(1 + K)}{1 - D_u(v)D_v(u)} = \frac{\delta}{1 - D_u(v)D_v(u)}.$$

Therefore, $1 - D_u(v)D_v(u) = \delta$, and so both $D_u(v) \geq 1 - \delta$ and $D_v(u) \geq 1 - \delta$. It follows that u and v are sufficiently close.

The conclusion of the proof follows the lines of the discussion in Section 2.7 (the proof of Lemma 2.4). Namely, the above analysis holds at every vertex i , and this contradicts the existence of a large exterior angle in an auxiliary Euclidean metric. \blacksquare

We are now equipped to prove Proposition 2.3.

2.9 The proof of Proposition 2.3

Proof of Proposition 2.3. The first two items follow immediately from Lemma 2.4. The argument for the third item is similar. In particular, we apply the same logic of Lemma 2.4 to the length function Λ , instead of the function F , to obtain a similar statement: if one edge of an r -periodic Finsler billiard trajectory is short, then so are all its edges. Therefore there exists some $\varepsilon > 0$ such that all critical points of Λ are contained in the interior of $G_\varepsilon(M, r)$.

Let η be a fixed number satisfying Lemma 2.16, and let R and ρ be the constants from Lemma 2.6. To satisfy the fourth item of the proposition, choose ε small enough so that $\eta > \frac{A}{\rho}\varepsilon^{\frac{1}{2r}}$, where $A > 0$ is a constant we will specify shortly. Assume that all lengths ℓ_i satisfy $\ell_i < \eta\rho$. Then, by Lemma 2.6, all corresponding u_i and v_i are η -tangent, so the conditions of Lemma 2.16 are satisfied, confirming the existence of an appropriate vector ν . Otherwise, there exists an index j such that $\ell_j \geq \eta\rho > A\varepsilon^{\frac{1}{2r}}$. Using the fact that the product of the squared lengths is ε , we write

$$\varepsilon^{\frac{1}{2}} = \prod_{1 \leq i \leq r} \ell_i > A\varepsilon^{\frac{1}{2r}} \prod_{i \neq j} \ell_i.$$

Therefore, there exists an index k such that $\ell_k < \varepsilon^{\frac{1}{2r}} A^{-\frac{1}{r-1}}$. It follows that

$$A^{\frac{r}{r-1}} < \frac{\ell_j}{\ell_k} = \frac{\ell_j}{\ell_{j-1}} \cdot \frac{\ell_{j-1}}{\ell_{j-2}} \cdots \frac{\ell_{k+1}}{\ell_k},$$

where the indices are understood cyclically. There are at most $r - 1$ factors on the right side, therefore, there exists some index i such that

$$A^{\frac{r}{(r-1)^2}} < \frac{\ell_{i+1}}{\ell_i}.$$

Now let $A = \left(\frac{R}{\rho}\right)^{\frac{(r-1)^2}{r}} \cdot C^{\frac{(r-1)^2}{mr}}$, where C and m are the constants determined in Lemma 2.14. Then we have

$$\frac{R}{\rho} \cdot C^{\frac{1}{m}} < \frac{\ell_{i+1}}{\ell_i} \leq -\frac{R}{\rho} \frac{p_i(v_i)}{p_i(u_i)},$$

where the second inequality is Lemma 2.6. It follows that

$$1 < \frac{p_i(v_i)^{2m}}{C^2 p_i(u_i)^{2m}}. \quad (2.9)$$

Assume there is no vector $\nu \in T_{x_i}M$ such that $D_{u_i}(\nu) > 0$ and $D_{v_i}(\nu) < 0$, since otherwise, we are done. Then the hypotheses of Corollary 2.15 hold, and we write

$$\frac{1 - D_{v_i}(w)}{1 - D_{u_i}(w)} \geq \frac{p_i(v_i)^{2m}}{C^2 p_i(u_i)^{2m}} > 1,$$

where the second inequality is (2.9). Therefore $D_{u_i}(w) > D_{v_i}(w)$, and since $\ell_{i+1} > \ell_i$, we also have $D_{u_i}(w)\ell_{i+1} > D_{v_i}(w)\ell_i$. Thus the vector $V \in T_x G_\varepsilon(M, r)$, whose only nonzero component is w , satisfies $d\Lambda(V) > 0$ and $dF(V) > 0$, as desired. \blacksquare

3 Topology of cyclic configuration spaces

Let $r \geq 2$ be an integer, and let M be a topological space. The *ordered configuration space* of r pairwise distinct points on M is the space

$$F(M, r) := \{(x_1, \dots, x_r) \in M^r : x_i \neq x_j \text{ for all } i \neq j\}.$$

The symmetric group on r letters \mathfrak{S}_r acts (from the left) on $F(M, r)$ by permuting the points, that is, for a permutation $\pi \in \mathfrak{S}_r$

$$\pi \cdot (x_1, \dots, x_r) = (x_{\pi(1)}, \dots, x_{\pi(r)}).$$

The *unlabeled configuration space* of r pairwise distinct points on M is the orbit space $F(M, r)/\mathfrak{S}_r$. We refer to F. Cohen [16] and Fadell & Husseini [19] for background on configuration spaces.

Let us repeat that the *ordered cyclic configuration space* of $r \geq 2$ consecutively distinct points on M is the space

$$G(M, r) := \{(x_1, \dots, x_r) \in M^r : x_i \neq x_{i+1} \text{ for all } i\},$$

where by convention $x_{r+1} = x_1$. Clearly, $F(M, r) \subseteq G(M, r) \subset M^r$. The Dihedral group $D_r = \langle a, b : a^r = b^2 = 1, ab = ba^{r-1} \rangle$ acts naturally on $G(M, r)$ by

$$\begin{aligned} a \cdot (x_1, x_2, \dots, x_{r-1}, x_r) &= (x_r, x_1, \dots, x_{r-2}, x_{r-1}), \\ b \cdot (x_1, x_2, \dots, x_{r-1}, x_r) &= (x_r, x_{r-1}, \dots, x_2, x_1). \end{aligned}$$

On the other hand, due to the geometric restriction coming from the Finsler distance being not symmetric, we consider only the action of the cyclic subgroup $\mathbb{Z}_r = \langle a \rangle$ on $G(M, r)$. Thus, in this paper, the *unlabeled cyclic configuration space* of r consecutively distinct points on M is the orbit space $G(M, r)/\mathbb{Z}_r$.

In this section we study the topology of the unlabeled cyclic configuration space $G(S^{d-1}, r)/\mathbb{Z}_r$ for r a prime. First, using an appropriate spectral sequence, we determine the cohomology of the unlabeled configuration space with coefficients in the field \mathbb{F}_r .

Theorem 3.1. *Let $r \geq 3$ be a prime.*

(1) *Let d be an even integer, and let*

$$\begin{aligned} A &:= \{\ell(d-2), \ell(d-2)+1 : 1 \leq \ell \leq r-2\}, \\ B &:= \{0, 1, \dots, (r-1)(d-2)+1\} \setminus A. \end{aligned}$$

Then

$$H^n(\mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r; \mathbb{F}_r) = \begin{cases} \mathbb{F}_r \oplus \mathbb{F}_r, & n \in A, \\ \mathbb{F}_r, & n \in B, \\ 0, & \text{otherwise.} \end{cases}$$

(2) *Let d be an odd integer, and let*

$$\begin{aligned} C &:= \{2\ell(d-2), 2\ell(d-2)+1 : 1 \leq \ell \leq \frac{r-3}{2}\}, \\ D &:= \{0, 1, \dots, (r-1)(d-2)+1\} \setminus C. \end{aligned}$$

Then

$$H^n(\mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r; \mathbb{F}_r) = \begin{cases} \mathbb{F}_r \oplus \mathbb{F}_r, & n \in C, \\ \mathbb{F}_r, & n \in D, \\ 0, & \text{otherwise.} \end{cases}$$

Second, using the same spectral sequence we derive the following estimate of the Lusternik–Schnirelmann category, which can be also deduced from the work of Karasev [29, Theorem 7].

Theorem 3.2. *Let $d \geq 3$ be an integer, and let $r \geq 3$ be a prime. Then the Lusternik–Schnirelmann category of the unlabeled cyclic configuration space $\mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r$ is bounded from below as follows:*

$$\mathrm{cat}(\mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r) \geq (r-1)(d-2)+1. \quad (3.1)$$

The estimate of the Lusternik–Schnirelmann category (3.1) of the unlabeled cyclic configuration space $\mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r$, in the case when r is a prime, is obtained by exhibiting an element of the cohomology $H^{(r-1)(d-2)+1}(\mathrm{B}(\mathbb{Z}_r); \mathbb{F}_r)$ that does not vanish along the homomorphism

$$H^*(\mathrm{B}(\mathbb{Z}_r); \mathbb{F}_r) \xrightarrow{p^*} H^*(\mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r),$$

which is induced by the unique, up to a homotopy, map $p: \mathrm{G}(S^{d-1}, r)/\mathbb{Z}_r \rightarrow \mathrm{B}(\mathbb{Z}_r)$, see Section 3.4. For this we substitute the map p with the projection map π_G of the fibration

$$\mathrm{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathrm{G}(S^{d-1}, r) \xrightarrow{\pi_G} \mathrm{B}(\mathbb{Z}_r)$$

and study the correspond Serre spectral sequence $E_*^{*,*}(\mathrm{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathrm{G}(S^{d-1}, r))$, see Section 3.1. The differentials of this spectral sequence are analyzed in Section 3.2 using the comparison of spectral sequences induced by the morphism of fiber bundles

$$\begin{array}{ccc} \mathrm{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathrm{F}(\mathbb{R}^{d-1}, r) & \xrightarrow{\mathrm{id} \times_{\mathbb{Z}_r} i} & \mathrm{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathrm{G}(S^{d-1}, r) \\ \downarrow \pi_F & & \downarrow \pi_G \\ \mathrm{B}(\mathbb{Z}_r) & \xrightarrow{\mathrm{id}} & \mathrm{B}(\mathbb{Z}_r). \end{array}$$

In this way we identify an element of the cohomology $u \in H^{(r-1)(d-2)+1}(\mathbb{Z}_r; \mathbb{F}_r)$ with the property that $\pi_G^*(u) \neq 0$, and consequently $p^*(u) \neq 0$. Using the classical concept of category weight of an element of cohomology, and its properties summarized in Lemma 3.10, we establish the lower bound (3.1).

3.1 Setting up a spectral sequence

Now we set up a spectral sequence that converges to the cohomology of the unlabeled cyclic configuration space $H^*(G(S^{d-1}, r)/\mathbb{Z}_r; \mathbb{F}_r)$. We use the fact that the group \mathbb{Z}_r acts freely on the cyclic configuration space $G(S^{d-1}, r)$.

Thus, since \mathbb{Z}_r acts freely on $G(S^{d-1}, r)$, the Borel construction $E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)$ and the quotient space $G(S^{d-1}, r)/\mathbb{Z}_r$ are homotopy equivalent. Indeed, the map

$$E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r) \longrightarrow G(S^{d-1}, r)/\mathbb{Z}_r,$$

induced by the \mathbb{Z}_r -equivariant projection on the second factor $E(\mathbb{Z}_r) \times G(S^{d-1}, r) \longrightarrow G(S^{d-1}, r)$, is a fibration with a contractible fiber $E(\mathbb{Z}_r)$, and therefore a homotopy equivalence. Hence,

$$H^*(G(S^{d-1}, r)/\mathbb{Z}_r; \mathbb{F}_r) \cong H^*(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r); \mathbb{F}_r),$$

and so we compute the cohomology of the Borel construction instead. An advantage of the Borel construction is that it is the total space of the following fibration

$$G(S^{d-1}, r) \longrightarrow E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r) \xrightarrow{\pi_G} B(\mathbb{Z}_r), \quad (3.2)$$

where π_G is induced by the \mathbb{Z}_r -equivariant projection on the first factor $E(\mathbb{Z}_r) \times G(S^{d-1}, r) \longrightarrow E(\mathbb{Z}_r)$.

The Serre spectral sequence induced by the fibration (3.2) has the E_2 -term given by

$$\begin{aligned} E_2^{p,q}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) &= \mathcal{H}^p(B(\mathbb{Z}_r); H^q(G(S^{d-1}, r); \mathbb{F}_r)) \\ &\cong H^p(\mathbb{Z}_r; H^q(G(S^{d-1}, r); \mathbb{F}_r)). \end{aligned} \quad (3.3)$$

Here $\mathcal{H}^*(\cdot)$ indicates that we have the cohomology with local coefficients; for more details consult for example [28, Section 3.H]. The local system is determined by the action of the fundamental group of the base space $\pi_1(B(\mathbb{Z}_r)) \cong \mathbb{Z}_r$ on the cohomology of a fiber of the fibration (3.2). The second notation assumes that the coefficients we have for the group cohomology of \mathbb{Z}_r are given in the \mathbb{Z}_r -module $H^q(G(S^{d-1}, r); \mathbb{F}_r)$. In the case when \mathbb{F}_r is a trivial \mathbb{Z}_r -module we set $H^*(\mathbb{Z}_r; \mathbb{F}_r) = \mathbb{F}_r[t] \otimes \Lambda(e)$ where $\deg(t) = 2$, $\deg(e) = 1$, and $\Lambda(\cdot)$ denotes the exterior algebra.

All the cohomologies we work with are \mathbb{F}_r vector spaces and therefore the Serre spectral sequence induced by the fibration (3.2) converges to the cohomology $H^*(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r); \mathbb{F}_r)$ as a vector space, that is

$$H^n(G(S^{d-1}, r)/\mathbb{Z}_r; \mathbb{F}_r) \cong H^n(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r); \mathbb{F}_r) \cong \bigoplus_{p+q=n} E_\infty^{p,q},$$

for every integer $n \geq 0$. In turn, since $G(S^{d-1}, r)/\mathbb{Z}_r$ is an open $(r(d-1))$ -manifold and has no cohomology in dimensions $\geq r(d-1)$, we have that $E_\infty^{p,q} = 0$ for all $p+q \geq r(d-1)$. For more details about Serre spectral sequences consult for example [34, Chapters 5–6] or [24, Chapter 3].

The first ingredient in the computation of the E_2 -term of the spectral sequence (3.3) is a description of the cohomology of the fiber $H^q(G(S^{d-1}, r); \mathbb{F}_r)$. For that we use the results of Farber [20, Theorems 18 and 19].

Proposition 3.3. *Let $d \geq 3$ be an integer, and let r be an odd prime.*

- (1) *If d is even, then the cohomology ring $H^*(G(S^{d-1}, r); \mathbb{F}_r)$ of the ordered cyclic configuration space $G(S^{d-1}, r)$ is generated by the elements*

$$\alpha, \beta_1, \beta_2, \dots, \beta_{r-2},$$

of degrees

$$\begin{aligned} \deg(\alpha) &= d - 1, \quad \deg(\beta_1) = d - 2, \quad \deg(\beta_2) = 2(d - 2), \quad \dots, \\ \deg(\beta_{r-2}) &= (r - 2)(d - 2), \end{aligned}$$

subject to the relations

$$\alpha^2 = 0, \quad \beta_i \beta_j = \begin{cases} \frac{(i+j)!}{i!j!} \beta_{i+j}, & \text{for } i+j \leq r-2, \\ 0, & \text{otherwise,} \end{cases}$$

for $1 \leq i \leq j \leq r-2$. In particular, for every $1 \leq k \leq r-2$ we have that $\beta_k = a_k \cdot \beta_1^k$ where $a_k \in \mathbb{F}_r \setminus \{0\}$.

- (2) If d is odd, then the cohomology ring $H^*(G(S^{d-1}, r); \mathbb{F}_r)$ of the ordered cyclic configuration space $G(S^{d-1}, r)$ is generated by the elements

$$\gamma, \delta_1, \delta_2, \dots, \delta_{\frac{r-3}{2}},$$

of degrees

$$\begin{aligned} \deg(\gamma) &= 2(d-2) + 1, \quad \deg(\delta_1) = 1 \cdot 2(d-2), \quad \deg(\delta_2) = 2 \cdot 2(d-2), \quad \dots, \\ \deg(\delta_{\frac{r-3}{2}}) &= \frac{r-3}{2} \cdot 2(d-2), \end{aligned}$$

subject to the relations

$$\gamma^2 = 0, \quad \delta_i \delta_j = \begin{cases} \frac{(i+j)!}{i!j!} \delta_{i+j}, & \text{for } i+j \leq \frac{r-3}{2}, \\ 0, & \text{otherwise,} \end{cases}$$

for $1 \leq i \leq j \leq \frac{r-3}{2}$. In particular, for every $1 \leq k \leq \frac{r-3}{2}$ we have that $\delta_k = b_k \cdot \delta_1^k$ where $b_k \in \mathbb{F}_r \setminus \{0\}$.

Transforming the previous information on the cohomology ring of the ordered cyclic configuration space $G(S^{d-1}, r)$ into an additive language we get the following corollary.

Corollary 3.4. *Let $d \geq 3$ be an integer, and let r be an odd prime.*

- (1) *If d is even and $A = \{\ell(d-2), \ell(d-2) + 1 : 1 \leq \ell \leq r-2\} \cup \{0, (r-1)(d-2) + 1\}$, then*

$$H^n(G(S^{d-1}, r); \mathbb{F}_r) = \begin{cases} \mathbb{F}_r, & n \in A, \\ 0, & \text{otherwise.} \end{cases}$$

- (2) *If d is odd and $B = \{2\ell(d-2), 2\ell(d-2) + 1 : 1 \leq \ell \leq \frac{r-3}{2}\} \cup \{0, (r-1)(d-2) + 1\}$, then*

$$H^n(G(S^{d-1}, r); \mathbb{F}_r) = \begin{cases} \mathbb{F}_r, & n \in B, \\ 0, & \text{otherwise.} \end{cases}$$

From Proposition 3.3 and its consequence Corollary 3.4 we can detect the action of the fundamental group of the base space $\pi_1(B(\mathbb{Z}_r)) \cong \mathbb{Z}_r$ on the cohomology of the ordered cyclic configuration space and compute the E_2 -term of the Serre spectral sequence (3.3); for an illustration of the E_2 -term in the case of even d see Fig. 5.

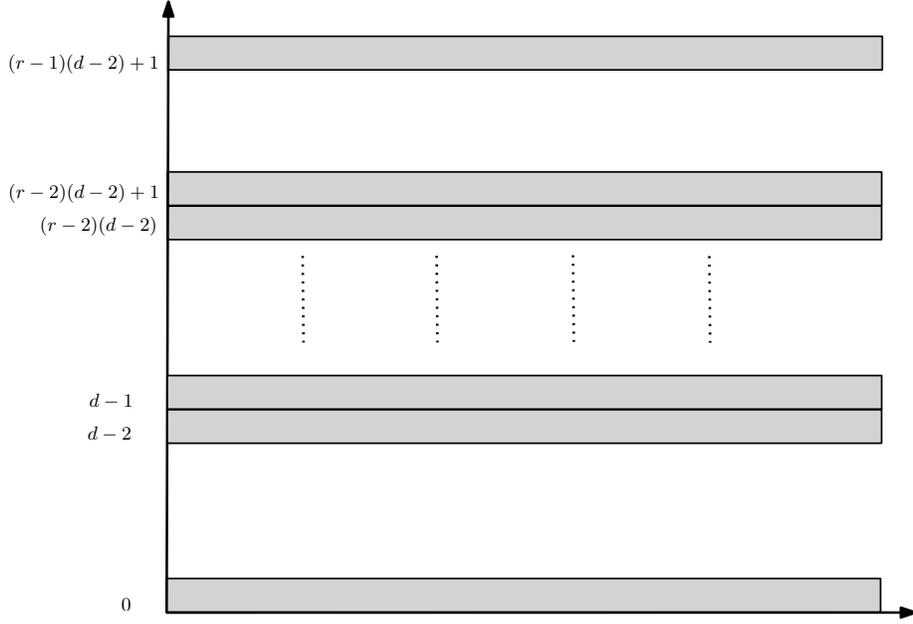


Figure 5. The E_2 -term of the Serre spectral sequence (3.3) in the case when $d \geq 4$ is even.

Corollary 3.5. *Let $d \geq 3$ be an integer, and let r be an odd prime. The E_2 -term of the Serre spectral sequence (3.3) has the following description.*

- (1) *If d is even and $A = \{\ell(d-2), \ell(d-2)+1 : 1 \leq \ell \leq r-2\} \cup \{0, (r-1)(d-2)+1\}$, then*

$$\begin{aligned} E_2^{p,q} &\cong H^p(\mathbb{Z}_r; H^q(G(S^{d-1}, r); \mathbb{F}_r)) \\ &\cong H^p(\mathbb{Z}_r; \mathbb{F}_r) \otimes H^q(G(S^{d-1}, r); \mathbb{F}_r) \cong \begin{cases} H^p(\mathbb{Z}_r; \mathbb{F}_r), & n \in A, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

- (2) *If d is odd and $B = \{2\ell(d-2), 2\ell(d-2)+1 : 1 \leq \ell \leq \frac{r-3}{2}\} \cup \{0, (r-1)(d-2)+1\}$, then*

$$\begin{aligned} E_2^{p,q} &\cong H^p(\mathbb{Z}_r; H^q(G(S^{d-1}, r); \mathbb{F}_r)) \\ &\cong H^p(\mathbb{Z}_r; \mathbb{F}_r) \otimes H^q(G(S^{d-1}, r); \mathbb{F}_r) \cong \begin{cases} H^p(\mathbb{Z}_r; \mathbb{F}_r), & n \in B, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Proof. In the fibration (3.2), the cohomology groups of the fiber $H^q(G(S^{d-1}, r); \mathbb{F}_r)$ are isomorphic either to 0 or to \mathbb{F}_r . Thus the group $\pi_1(B(\mathbb{Z}_r)) \cong \mathbb{Z}_r$ can only act trivially on $H^q(G(S^{d-1}, r); \mathbb{F}_r)$: Indeed, if $f: \mathbb{F}_r \rightarrow \mathbb{F}_r$ is an \mathbb{F}_r -linear map induced by the action of a generator of \mathbb{Z}_r , and $f(1) = a \in \mathbb{F}_r$, then by Fermat's little theorem $1 = \text{id}_{\mathbb{F}_r}(1) = f^r(1) = a^r = a$ and so $f = \text{id}_{\mathbb{F}_r}$. Consequently, the cohomology $H^q(G(S^{d-1}, r); \mathbb{F}_r)$ is a trivial \mathbb{Z}_r -module for every q , and the statement of the corollary follows. \blacksquare

Remark 3.6. It is important to point out that the previous corollary also implies that the differentials of the Serre spectral sequence (3.3) are not only $H^*(\mathbb{Z}_r; \mathbb{F}_r)$ -module morphisms, but even more, they satisfy the Leibniz rule; see [34, Definition 1.6 and Proposition 5.6].

3.2 Computing the spectral sequence

In this section, before making explicit computation, we consider the ordered configuration space $F(\mathbb{R}^{d-1}, r)$ as a subset of the ordered cyclic configuration space $G(S^{d-1}, r)$. The \mathbb{Z}_r -equivariant

inclusion $i: F(\mathbb{R}^{d-1}, r) \rightarrow G(S^{d-1}, r)$ induces the following morphism of Borel construction fibrations:

$$\begin{array}{ccc} E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r) & \xrightarrow{\text{id} \times_{\mathbb{Z}_r} i} & E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r) \\ \downarrow \pi_F & & \downarrow \pi_G \\ B(\mathbb{Z}_r) & \xrightarrow{\text{id}} & B(\mathbb{Z}_r). \end{array} \quad (3.4)$$

The morphism of fibrations (3.4) induces a morphism of associated Serre spectral sequences

$$E_*^{*,*}(\text{id} \times_{\mathbb{Z}_r} i): E_*^{*,*}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) \leftarrow E_*^{*,*}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)), \quad (3.5)$$

which is the identity homomorphism on the zero row of the E_2 -term, that is $E_2^{*,0}(\text{id} \times_{\mathbb{Z}_r} i) = \text{id}$.

The Serre spectral sequence $E_*^{*,*}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r))$ has been completely described first by F. Cohen in his seminal paper [16, Sections 8–10], and much later, from the point of view of the equivariant Goresky–MacPherson formula, in [12, Theorem 6.1]. For an illustration of the E_2 -term of this spectral sequence see Fig. 6. In particular, we will use the following facts proved in [12, 16].

Proposition 3.7. *Let $d \geq 3$ be an integer, and let r be an odd prime. The Serre spectral sequence associated to the Borel construction fibration*

$$F(\mathbb{R}^{d-1}, r) \longrightarrow E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r) \xrightarrow{\pi_F} B(\mathbb{Z}_r), \quad (3.6)$$

with the E_2 -term

$$E_2^{p,q} = \mathcal{H}^p(B(\mathbb{Z}_r); H^q(F(\mathbb{R}^{d-1}, r); \mathbb{F}_r)) \cong H^p(\mathbb{Z}_r; H^q(F(\mathbb{R}^{d-1}, r); \mathbb{F}_r))$$

has to following properties:

- (1) For all $p \geq 1$ and $1 \leq q \leq (r-1)(d-2) - 1$, we have that

$$E_2^{p,q}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) = E_\infty^{p,q}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) = 0.$$

- (2) For all $p \in \mathbb{Z}$ and $q \geq (r-1)(d-2) + 1$, we have that

$$E_2^{p,q}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) = E_\infty^{p,q}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) = 0.$$

- (3) For all $p \geq 0$ and $2 \leq s \leq (r-1)(d-2)$, the differential

$$\partial_s: E_s^{p,s-1}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) \longrightarrow E_s^{p+s,0}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^d, r))$$

vanishes. Consequently, for $0 \leq p \leq (r-1)(d-2)$, we have that

$$H^p(\mathbb{Z}_r; \mathbb{F}_r) = E_2^{p,0}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) = E_\infty^{p,0}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)).$$

- (4) The only non-zero differential of the spectral sequence is

$$\begin{aligned} \partial_{(d-2)(r-1)+1}: E_{(d-2)(r-1)+1}^{p,(d-2)(r-1)}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) \\ \longrightarrow E_{(d-2)(r-1)+1}^{p+(d-2)(r-1)+1,0}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)), \end{aligned}$$

and is an isomorphism for all $p \geq 0$. Consequently, for all $p \geq (r-1)(d-2) + 1$, one has

$$E_{(d-2)(r-1)+2}^{p,0}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) \cong E_\infty^{p,0}(E(\mathbb{Z}_r) \times_{\mathbb{Z}_r} F(\mathbb{R}^{d-1}, r)) = 0.$$

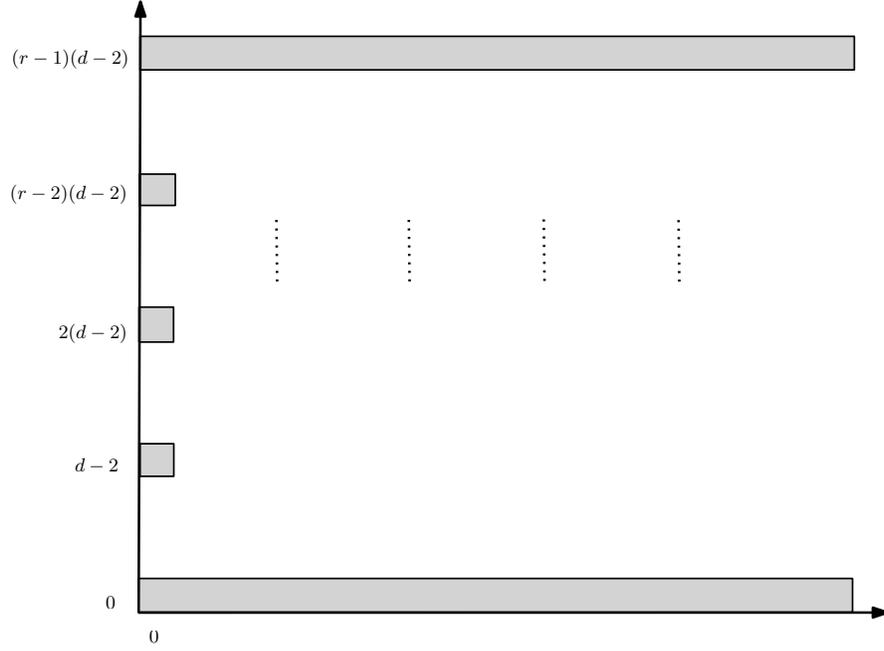


Figure 6. The $E_2 = E_{(r-1)(d-2)}$ -term of the Serre spectral sequence of the fibration (3.6).

From the existence of the morphism $E_*^{*,*}(\text{id} \times_{\mathbb{Z}_r} i)$ of the Serre spectral sequences (3.5), the fact that $E_2^{*,0}(\text{id} \times_{\mathbb{Z}_r} i)$ is the identity, and the following equality from Proposition 3.7:

$$E_\infty^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{F}(\mathbb{R}^{d-1}, r)) = E_2^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{F}(\mathbb{R}^{d-1}, r)) = H^p(\mathbb{Z}_r; \mathbb{F}_r),$$

which holds for $0 \leq p \leq (r-1)(d-2)$, we deduce an important property of the Serre spectral sequence $E_*^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r))$ stated in the next corollary.

Corollary 3.8. *Let $d \geq 3$ be an integer, and let r be an odd prime. For $0 \leq p \leq (r-1)(d-2)$, one has*

$$E_\infty^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong E_2^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong H^p(\mathbb{Z}_r; \mathbb{F}_r).$$

In particular, all the differentials

$$\partial_s : E_s^{p,s-1}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \longrightarrow E_s^{p+s,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r))$$

vanish for $p \geq 0$ and $2 \leq s \leq (r-1)(d-2)$.

Combining the previous fact about the Serre spectral sequence $E_*^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r))$ with the observation that $E_\infty^{p,q}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) = 0$ for all $p+q \geq r(d-1)$, we have that for some $s \geq (r-1)(d-2) + 1$ the differential

$$\partial_s : E_s^{p,s-1}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \longrightarrow E_s^{p+s,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r))$$

does *not* vanish. In particular,

$$E_\infty^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) = 0$$

for all $p \geq r(d-1)$.

Now the properties of the Serre spectral sequence $E_*^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r))$, derived so far, will be used to completely compute it. The computation proceeds in two separate steps, depending on the parity of d .

3.2.1 d is an even integer, $d \geq 4$

The Serre spectral sequence $E_*^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r))$ has a multiplicative structure, and differentials satisfy the Leibniz rule. Thus, to compute the spectral sequence completely it suffices to determine values of the differentials only on the generators of the cohomology ring of the fiber. In this particular case, we need to evaluate the differentials on the elements $1 \otimes \alpha, 1 \otimes \beta_1, \dots, 1 \otimes \beta_{r-1}$. It is important to notice that

$$(1 \otimes \beta_1)^i = 1 \otimes \beta_1^i = a_i \cdot (1 \otimes \beta_i)$$

for $1 \leq i \leq r-1$ and some $a_i \in \mathbb{F}_r \setminus \{0\}$.

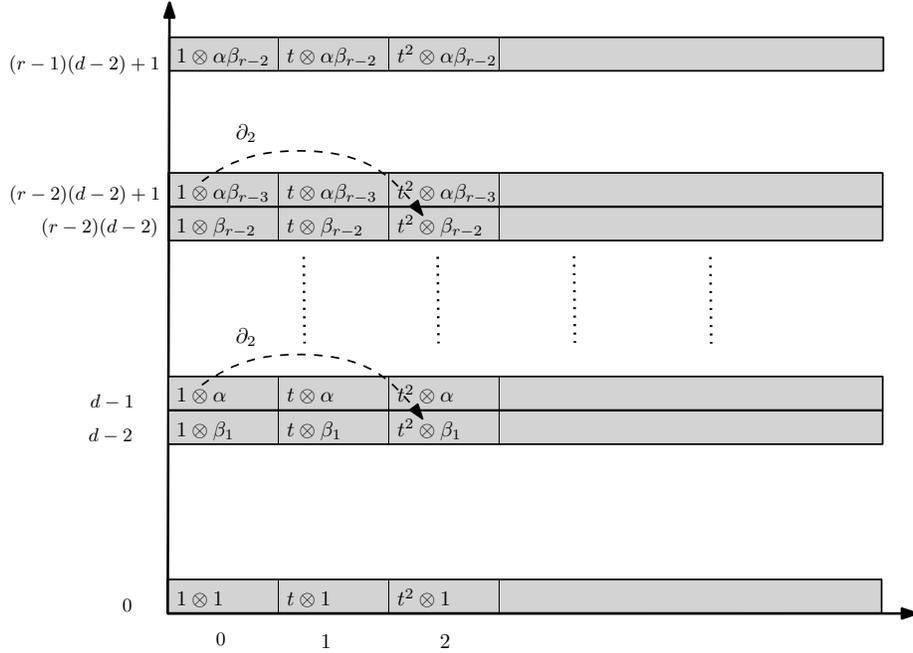


Figure 7. For $d \geq 4$ even the E_2 -term with differentials of the Serre spectral sequence (3.3).

Consider the differential ∂_2 . The degrees of the cohomology generators $\beta_1, \dots, \beta_{r-1}$ are $(d-2), \dots, (r-2)(d-2)$, respectively. Since $H^i(G(S^{d-1}, r); \mathbb{F}_r) = 0$ vanishes for all $i \in \{(d-2)-1, \dots, (r-2)(d-2)-1\}$, one has

$$\partial_2(1 \otimes \beta_1) = \partial_2(1 \otimes \beta_2) = \dots = \partial_2(1 \otimes \beta_{r-1}) = 0. \quad (3.7)$$

Thus, to completely determine the second differential we need to find the value

$$\partial_2(1 \otimes \alpha) \in E_2^{2, d-2}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)).$$

Let us assume that $\partial_2(1 \otimes \alpha) = 0$. Then, due to the multiplicative property of the Serre spectral sequence, we have that the second differential ∂_2 vanishes, implying that

$$E_2^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong E_3^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)).$$

The vanishing of cohomology groups $H^*(G(S^{d-1}, r); \mathbb{F}_r)$ in appropriate dimensions implies that the next possible non-zero differential is ∂_{d-1} . More precisely, $\partial_{d-1}(1 \otimes \beta_1)$ might be non-zero while we know that $\partial_{d-1}(1 \otimes \alpha) = 0$. Since $d-1 \leq (r-1)(d-2)$, using Corollary 3.8, we get that $\partial_{d-1}(1 \otimes \beta_1) = 0$, and consequently the differential ∂_{d-1} vanishes. For the next differential

we know that $\partial_d(1 \otimes \beta_1) = 0$, and since $d \leq (r-1)(d-2)$ from Corollary 3.8, we have that $\partial_d(1 \otimes \alpha) = 0$. Hence, the differential ∂_d also vanishes. In particular this means that

$$E_2^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong \cdots \cong E_{d+1}^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)).$$

Since for $s \geq d+1$ all differentials $\partial_s(1 \otimes \beta_1)$ and $\partial_s(1 \otimes \alpha)$ are zero, the multiplicative property of the Serre spectral sequence implies that all the differentials ∂_s vanish. Thus

$$E_2^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong E_\infty^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)).$$

This is a contradiction to the fact, observed after Corollary 3.8, that at least one of the differentials

$$\partial_s: E_s^{p,s-1}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \longrightarrow E_s^{p+s,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)),$$

for $s \geq (r-1)(d-2) + 1$, does not vanish. Therefore, $\partial_2(1 \otimes \alpha) \neq 0$.

Set $\partial_2(1 \otimes \alpha) = a \cdot (t \otimes \beta_1)$ where $a \in \mathbb{F}_r \setminus \{0\}$; for illustration see Fig. 7. The multiplicative property of the Serre spectral sequence, combined with (3.7), yields that

$$\begin{aligned} \partial_2(1 \otimes \alpha \beta_i) &= \partial_2((1 \otimes \alpha)(1 \otimes \beta_i)) = \partial_2(1 \otimes \alpha)(1 \otimes \beta_i) + (-1)^{d-1}(1 \otimes \alpha)\partial_2(1 \otimes \beta_i) \\ &= \partial_2(1 \otimes \alpha)(1 \otimes \beta_i) = a \cdot (t \otimes \beta_1)(1 \otimes \beta_i) \\ &= \begin{cases} \left(a \frac{(i+1)!}{i!!} \right) \cdot (t \otimes \beta_{i+1}), & 1 \leq i \leq r-3, \\ 0, & i = r-2. \end{cases} \end{aligned}$$

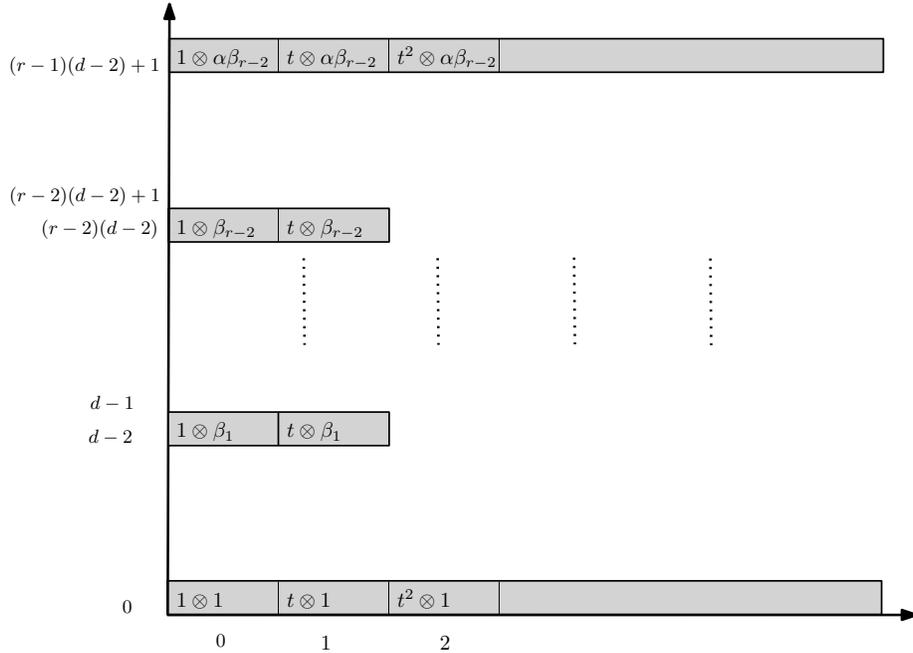


Figure 8. For $d \geq 4$ even the E_3 -term of the Serre spectral sequence (3.3).

Thus the E_3 -term, illustrated in Fig. 8, is given by

$$E_3^{p,q}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong \begin{cases} H^p(\mathbb{Z}_r; \mathbb{F}_r), & q \in \{0, (r-1)(d-2) + 1\}, \\ \mathbb{F}_r, & p \in \{0, 1\}, q \in \{(d-2), \dots, (r-2)(d-2)\}, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover the multiplicative property of the Serre spectral sequence implies that all the differentials $\partial_3, \dots, \partial_{(r-1)(d-2)+1}$ vanish. Therefore

$$E_3^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong \dots \cong E_{(r-1)(d-2)+2}^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)).$$

The differential $\partial_{(r-1)(d-2)+2}$ is the only remaining differential that can be non-zero. Since

$$E_2^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong \dots \cong E_{(r-1)(d-2)+2}^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)),$$

and we know that

$$E_2^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \neq 0 = E_{\infty}^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r))$$

for all $p \geq r(d-1)$, we conclude that $\partial_{(r-1)(d-2)+2} \neq 0$. The multiplicative property of the Serre spectral sequence again yields that $\partial_{(r-1)(d-2)+2}$ is completely determined by its image at $1 \otimes \alpha\beta_{r-2}$, which has to be non-zero. Set

$$\partial_{(r-1)(d-2)+2}(1 \otimes \alpha\beta_{r-2}) = b \cdot (t^{\frac{(r-1)(d-2)+2}{2}} \otimes 1)$$

for some $b \in \mathbb{F}_r \setminus \{0\}$, as illustrated in Fig. 9. Consequently, the differentials

$$\begin{aligned} \partial_{(r-1)(d-2)+2}: E_s^{p,(r-1)(d-2)+1}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \\ \longrightarrow E_s^{p+(r-1)(d-2)+2,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \end{aligned}$$

are isomorphisms for all $p \geq 0$.

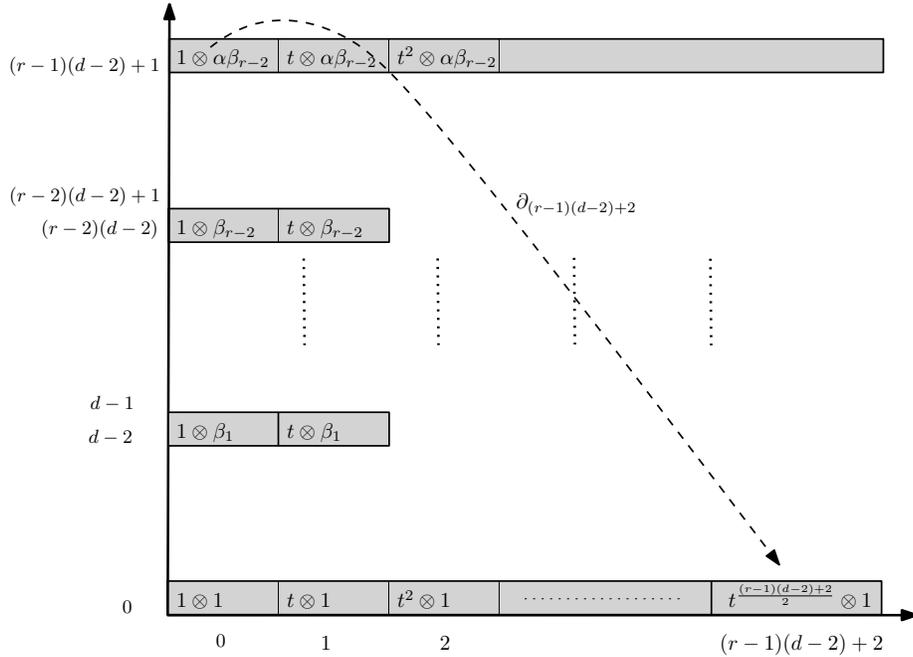


Figure 9. For $d \geq 4$ even the differential $\partial_{(r-1)(d-2)+2}$ of the Serre spectral sequence (3.3).

Hence, the E_{∞} -term is given by

$$E_{\infty}^{p,q}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbb{G}(S^{d-1}, r)) \cong \begin{cases} H^p(\mathbb{Z}_r; \mathbb{F}_r), & q = 0, 0 \leq p \leq (r-1)(d-2) + 1, \\ \mathbb{F}_r, & p \in \{0, 1\}, q \in \{(d-2), \dots, (r-2)(d-2)\}, \\ 0, & \text{otherwise,} \end{cases} \quad (3.8)$$

and illustrated in Fig. 10. Moreover, we have obtained that the map π_G^* induced by the projection map of the fibration (3.2) in cohomology

$$\pi_G^*: H^p(\mathbb{Z}_r; \mathbb{F}_r) \longrightarrow H^p(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r); \mathbb{F}_r) \quad (3.9)$$

is an injection for $0 \leq p \leq (r-1)(d-2) + 1$.

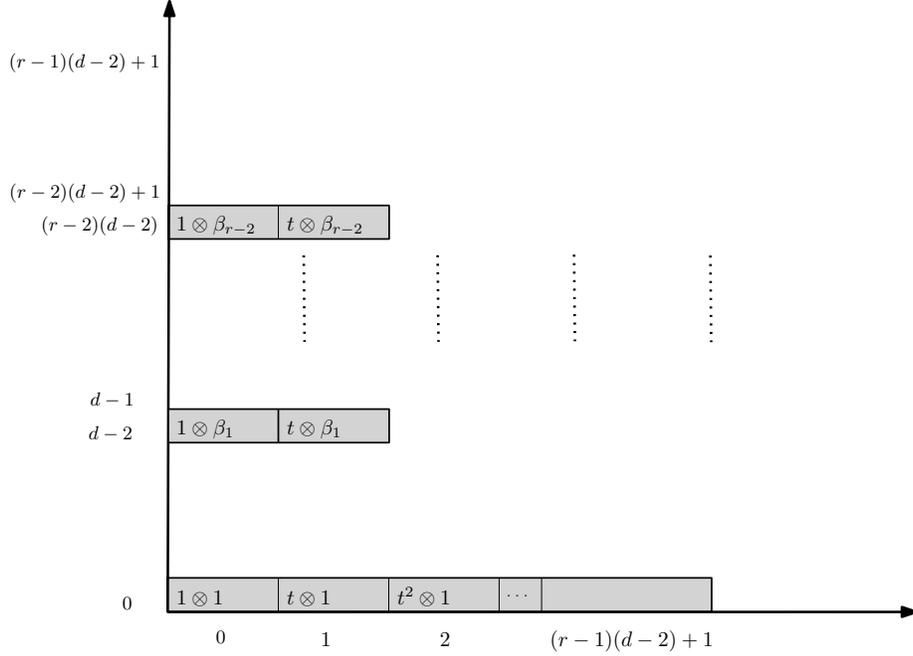


Figure 10. For $d \geq 4$ even the E_∞ -term of the Serre spectral sequence (3.3).

3.2.2 d is an odd integer, $d \geq 3$

The Serre spectral sequence $E_*^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r))$ has a multiplicative structure. Hence, to compute it completely, we will determine values of the differentials on the generators of the cohomology ring of the fiber. In this particular case, we need to evaluate the differentials on the elements $\gamma, \delta_1, \dots, \delta_{\frac{r-3}{2}}$. Observe that

$$(1 \otimes \delta_1)^i = 1 \otimes \delta_1^i = a_i \cdot (1 \otimes \delta_i)$$

for $1 \leq i \leq \frac{r-3}{2}$ and some $a_i \in \mathbb{F}_r \setminus \{0\}$.

Similarly to Section 3.2.1, we consider first the values of the differential ∂_2 on the generators. Since the generator δ_1 is of degree 2 and $H^1(G(S^{d-1}, r); \mathbb{F}_r) = 0$ we have that $\partial_2(1 \otimes \delta_1) = 0$. Further, the Leibniz rule implies that

$$\partial_2(1 \otimes \delta_i) = a_i \cdot \partial_2((1 \otimes \delta_1)^i) = (ia_i) \cdot \partial_2(1 \otimes \delta_1) = 0$$

for $1 \leq i \leq \frac{r-3}{2}$ and some $a_i \in \mathbb{F}_r \setminus \{0\}$. Thus, in order to determine the second differential ∂_2 , we need to determine

$$\partial_2(1 \otimes \gamma) \in E_2^{2, 2(d-2)}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)).$$

Assume that $\partial_2(1 \otimes \gamma) = 0$. Then the multiplicative property of the Serre spectral sequence implies that the second differential ∂_2 gives

$$E_2^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong E_3^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)).$$

The vanishing of cohomology groups $H^*(G(S^{d-1}, r); \mathbb{F}_r)$ in appropriate dimensions implies that the next possible non-zero differential is $\partial_{2(d-2)+1}$. More precisely, $\partial_{2(d-2)+1}(1 \otimes \delta_1)$ might be non-zero, while we know that $\partial_{2(d-2)+1}(1 \otimes \gamma) = 0$. Since $2(d-2)+1 \leq (r-1)(d-2)$ from Corollary 3.8, we have that $\partial_{2(d-2)+1}(1 \otimes \delta_1) = 0$. Consequently, the differential $\partial_{2(d-2)+1}$ vanishes. About the next differential we know that $\partial_{2(d-2)+2}(1 \otimes \delta_1) = 0$. Since $2(d-2)+2 \leq (r-1)(d-2)$, again Corollary 3.8 implies that $\partial_{2(d-2)+2}(1 \otimes \gamma) = 0$. Therefore, the differential $\partial_{2(d-2)+2}$ also vanishes. This means that

$$E_2^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong \dots \cong E_{2(d-2)+3}^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)).$$

For $s \geq 2(d-2)+3$, all differentials $\partial_s(1 \otimes \delta_1)$ and $\partial_s(1 \otimes \gamma)$ are zero. Hence, the multiplicative property of the Serre spectral sequence implies that, for $s \geq 2(d-2)+3$, all the differentials ∂_s vanish. Thus

$$E_2^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong E_\infty^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)),$$

which is a contradiction to the fact, observed after Corollary 3.8, that at least one of the differentials

$$\partial_s: E_s^{p,s-1}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \longrightarrow E_s^{p+s,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)),$$

for $s \geq (r-1)(d-2)+1$, does not vanish. Hence, $\partial_2(1 \otimes \gamma) \neq 0$.

Now set $\partial_2(1 \otimes \gamma) = a \cdot (t \otimes \delta_1)$ where $a \in \mathbb{F}_r \setminus \{0\}$. The multiplicative property yields

$$\begin{aligned} \partial_2(1 \otimes \gamma \delta_i) &= \partial_2((1 \otimes \gamma)(1 \otimes \delta_i)) = \partial_2(1 \otimes \gamma)(1 \otimes \delta_i) + (-1)^{d-1}(1 \otimes \gamma)\partial_2(1 \otimes \delta_i) \\ &= \partial_2(1 \otimes \gamma)(1 \otimes \delta_i) = a \cdot (t \otimes \gamma_1)(1 \otimes \gamma_i) \\ &= \begin{cases} \left(a \frac{(i+1)!}{i!1!}\right) \cdot (t \otimes \gamma_{i+1}), & 1 \leq i \leq \frac{r-3}{2}, \\ 0, & i = r-2. \end{cases} \end{aligned}$$

Consequently, the E_3 -term is

$$E_3^{p,q}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong \begin{cases} H^p(\mathbb{Z}_r; \mathbb{F}_r), & q \in \{0, (r-1)(d-2)+1\}, \\ \mathbb{F}_r, & p \in \{0, 1\}, q \in \{2(d-2), \dots, \frac{(r-3)}{2}2(d-2)\}, \\ 0, & \text{otherwise.} \end{cases}$$

Further, the multiplicative property implies that all the differentials $\partial_3, \dots, \partial_{(r-1)(d-2)+1}$ vanish. Thus,

$$E_3^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong \dots \cong E_{(r-1)(d-2)+2}^{*,*}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)).$$

The differential $\partial_{(r-1)(d-2)+2}$ is the last remaining differential that can be non-zero. Since for all p

$$E_2^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \cong \dots \cong E_{(r-1)(d-2)+2}^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)),$$

and we know that

$$E_2^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r)) \neq 0 = E_\infty^{p,0}(\mathbb{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r))$$

for $p \geq r(d-1)$, we have that indeed $\partial_{(r-1)(d-2)+2} \neq 0$. The multiplicative property again yields that $\partial_{(r-1)(d-2)+2}$ is determined by its image at $1 \otimes \gamma \delta_{\frac{r-3}{2}}$ and has to be non-zero. Set

$$\partial_{(r-1)(d-2)+2}(1 \otimes \gamma \delta_{\frac{r-3}{2}}) = b \cdot \left(t^{\frac{(r-1)(d-2)+2}{2}} \otimes 1\right)$$

for some $b \in \mathbb{F}_r \setminus \{0\}$. Thus, the homomorphisms

$$\begin{aligned} \partial_{(r-1)(d-2)+2}: E_s^{p,(r-1)(d-2)+1}(\mathbf{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbf{G}(S^{d-1}, r)) \\ \longrightarrow E_s^{p+(r-1)(d-2)+2,0}(\mathbf{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbf{G}(S^{d-1}, r)) \end{aligned}$$

are isomorphism for all $p \geq 0$. Hence, the E_∞ -term is given by

$$\begin{aligned} E_\infty^{p,q}(\mathbf{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbf{G}(S^{d-1}, r)) \\ \cong \begin{cases} H^p(\mathbb{Z}_r; \mathbb{F}_r), & q = 0, 0 \leq p \leq (r-1)(d-2) + 1, \\ \mathbb{F}_r, & p \in \{0, 1\}, q \in \{2(d-2), \dots, \frac{(r-3)}{2}2(d-2)\}, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (3.10)$$

Furthermore, we have obtained again that the map π_G^* , induced by the projection map of the fibration (3.2) in cohomology

$$\pi_G^*: H^p(\mathbb{Z}_r; \mathbb{F}_r) \longrightarrow H^p(\mathbf{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbf{G}(S^{d-1}, r); \mathbb{F}_r), \quad (3.11)$$

is an injection for $0 \leq p \leq (r-1)(d-2) + 1$.

3.3 Proof of Theorem 3.1

Recall that for every $n \in \mathbb{N}$ we have isomorphisms of vector spaces

$$H^n(\mathbf{G}(S^{d-1}, r)/\mathbb{Z}_r; \mathbb{F}_r) \cong H^n(\mathbf{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} \mathbf{G}(S^{d-1}, r); \mathbb{F}_r) \cong \bigoplus_{p+q=n} E_\infty^{p,q}.$$

Since $H^p(\mathbb{Z}_r; \mathbb{F}_r) = \mathbb{F}_r$ for all p , the relations (3.8) for the case of $d \geq 4$ even, and (3.10) for the case of $d \geq 3$ odd, imply the statement of the theorem.

3.4 Proof of Theorem 3.2

In order to use the results of Section 3.2 for a proof of an estimate on the Lusternik–Schnirelmann category of the unlabeled cyclic configuration space $\mathbf{G}(S^{d-1}, r)/\mathbb{Z}_r$, we first recall some basic notions and results concerning the Lusternik–Schnirelmann category.

The *Lusternik–Schnirelmann category* of a topological space X , denoted by $\text{cat}(X)$, is the smallest integer k for which the space X can be covered by $k+1$ open subsets U_1, U_2, \dots, U_{k+1} with the property that all inclusions $U_i \rightarrow X$ are nullhomotopic. Some key properties of the Lusternik–Schnirelmann category that we will use are given in the next lemma, see, e.g., [17].

Lemma 3.9.

- (1) If X is homotopy equivalent to Y , then $\text{cat}(X) = \text{cat}(Y)$.
- (2) If $p: X \rightarrow Y$ is a covering, then $\text{cat}(X) \leq \text{cat}(Y)$.
- (3) If X is a $(k-1)$ -connected CW-complex, then $\text{cat}(X) \leq \frac{1}{k} \dim(X)$.

Let X be a topological space, and let R be a commutative ring with unit. We need the notion of the *category weight* of an element $u \in H^*(X; R)$. Originally introduced by Fadell and Husseini [18], we use the homotopy invariant version of this notion due to Rudyak [37] and Strom [39]. See [17, Section 2.7, p. 62; Section 8.3, p. 240] for details.

In the next lemma we list properties of the category weight that we will use in the proof of Theorem 3.2. For more details on this lemma consult for example [17, Proposition 8.22, pp. 242–243, p. 259] and [36, Proposition 2.2(3)].

Lemma 3.10. *Let R be a commutative ring with unit.*

- (1) *If $0 \neq u \in H^\ell(X; R)$, then $\text{wgt}(u) \leq \text{cat}(X)$.*
- (2) *Let $f: X \rightarrow Y$ be a continuous map, and let $u \in H^\ell(Y; R)$. If $0 \neq f^*(u) \in H^\ell(X; R)$, then $\text{wgt}(u) \leq \text{wgt}(f^*(u))$.*
- (3) *If G is a finite group and $0 \neq u \in H^\ell(\text{BG}; R)$, then $\ell = \text{wgt}(u)$.*

In the following we combine previously established results to give a proof of Theorem 3.2. Recall that the group \mathbb{Z}_r acts freely on the cyclic configuration space $G(S^{d-1}, r)$ and consider the following diagram, which commutes up to a homotopy

$$\begin{array}{ccc} & \xrightarrow{\pi_G} & \\ \text{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r) & \xrightarrow{p} G(S^{d-1}, r)/\mathbb{Z}_r & \xrightarrow{c} \text{B}(\mathbb{Z}_r), \end{array} \quad (3.12)$$

where the map p is induced by the projection on the second factor $\text{E}(\mathbb{Z}_r) \times G(S^{d-1}, r) \rightarrow G(S^{d-1}, r)$ and is a homotopy equivalence, the map π_G is the projection in the Borel construction fibration (3.2) induced by the projection on the first factor $\text{E}(\mathbb{Z}_r) \times G(S^{d-1}, r) \rightarrow \text{E}(\mathbb{Z}_r)$, and c is a classifying map associated to the free \mathbb{Z}_r action on $G(S^{d-1}, r)$. Uniqueness of the classifying map up to a homotopy implies that the diagram (3.12) commutes up to a homotopy; for background see for example [1, Section II.1].

Applying the cohomology functor $H^*(\cdot; \mathbb{F}_r)$ to the diagram (3.12), we get the following diagram of abelian groups that commutes up to an isomorphism

$$\begin{array}{ccc} & \xrightarrow{\pi_G^*} & \\ H^*(\text{E}(\mathbb{Z}_r) \times_{\mathbb{Z}_r} G(S^{d-1}, r); \mathbb{F}_r) & \xleftarrow{p^*} H^*(G(S^{d-1}, r)/\mathbb{Z}_r) & \xleftarrow{c^*} H^*(\text{B}(\mathbb{Z}_r); \mathbb{F}_r) \\ & \cong H^*(\mathbb{Z}_r; \mathbb{F}_r). & \end{array} \quad (3.13)$$

Let

$$u := t^{\frac{(r-1)(d-2)}{2}} e \in H^{(r-1)(d-2)+1}(\mathbb{Z}_r; \mathbb{F}_r).$$

Then, according to (3.9) and (3.11), we have that $\pi_G^*(u) \neq 0$. Consequently, from diagram (3.13), we get $c^*(u) \neq 0$. Now Lemma 3.10 yields that

$$\begin{aligned} \text{wgt}(u) &= (r-1)(d-2) + 1, & \text{wgt}(u) &\leq \text{wgt}(c^*(u)), \\ \text{wgt}(c^*(u)) &\leq \text{cat}(G(S^{d-1}, r)/\mathbb{Z}_r). \end{aligned}$$

Hence,

$$\text{cat}(G(S^{d-1}, r)/\mathbb{Z}_r) \geq (r-1)(d-2) + 1,$$

and the proof of Theorem 3.2 is complete. ■

4 Proofs of the main results

According to Section 2.4, the number of \mathbb{Z}_r -orbits of the critical points of the length function on the cyclic configuration space $G(M, r)$ is bounded below by the category of the space $G(S^{d-1}, r)/\mathbb{Z}_r$ (the Lusternik–Schnirelmann theory) or, in general position, by the sum of Betti numbers of $G(S^{d-1}, r)/\mathbb{Z}_r$ (the Morse theory). Thus item (A) of Theorem 1.2 follows from Theorem 3.2, and item (B) follows from Theorem 3.1.

Acknowledgements

We are grateful to Sergei Ivanov for useful discussions on Finsler geometry, and we are grateful to the following sources of funding. Pavle V. M. Blagojević, Serge Tabachnikov, and Günter M. Ziegler were supported by the DFG via the Collaborative Research Center TRR 109 “Discretization in Geometry and Dynamics”. Pavle V. M. Blagojević was supported by the grant ON 174024 of Serbian Ministry of Education and Science. Michael Harrison and Serge Tabachnikov were supported by the NSF grant DMS-1510055. We are also grateful to the referees for their suggestions.

References

- [1] Adem A., Milgram R.J., Cohomology of finite groups, 2nd ed., *Grundlehren der Mathematischen Wissenschaften*, Vol. 309, [Springer-Verlag](#), Berlin, 2004.
- [2] Akopyan A.V., Zaslavsky A.A., Geometry of conics, *Mathematical World*, Vol. 26, [Amer. Math. Soc.](#), Providence, RI, 2007.
- [3] Álvarez J.C., Durán C., An introduction to Finsler geometry, *Notas de la Escuela Venezolana de Matemáticas*, 1998.
- [4] Artstein-Avidan S., Karasev R., Ostrover Y., From symplectic measurements to the Mahler conjecture, *Duke Math. J.* **163** (2014), 2003–2022, [arXiv:1303.4197](#).
- [5] Artstein-Avidan S., Ostrover Y., Bounds for Minkowski billiard trajectories in convex bodies, *Int. Math. Res. Not.* **2014** (2014), 165–193, [arXiv:1111.2353](#).
- [6] Babenko I.K., Periodic trajectories of three-dimensional Birkhoff billiards, *Math. USSR Sb.* **71** (1992), 1–13.
- [7] Bangert V., On the existence of closed geodesics on two-spheres, *Internat. J. Math.* **4** (1993), 1–10.
- [8] Bao D., Chern S.-S., Shen Z., An introduction to Riemann–Finsler geometry, *Graduate Texts in Mathematics*, Vol. 200, [Springer-Verlag](#), New York, 2000.
- [9] Berglund N., Kunz H., Integrability and ergodicity of classical billiards in a magnetic field, *J. Statist. Phys.* **83** (1996), 81–126, [arXiv:chao-dyn/9501009](#).
- [10] Bialy M., On totally integrable magnetic billiards on constant curvature surface, *Electron. Res. Announc. Math. Sci.* **19** (2012), 112–119, [arXiv:1208.2455](#).
- [11] Birkhoff G.D., On the periodic motions of dynamical systems, *Acta Math.* **50** (1927), 359–379.
- [12] Blagojević P.V.M., Lück W., Ziegler G.M., Equivariant topology of configuration spaces, *J. Topol.* **8** (2015), 414–456, [arXiv:1207.2852](#).
- [13] Busemann H., Problem IV: Desarguesian spaces, in *Mathematical Developments Arising from Hilbert Problems* (Proc. Sympos. Pure Math., Northern Illinois Univ., De Kalb, Ill., 1974), 1976, 131–141.
- [14] Chern S.-S., Finsler geometry is just Riemannian geometry without the quadratic restriction, *Notices Amer. Math. Soc.* **43** (1996), 959–963.
- [15] Chernov N., Markarian R., Chaotic billiards, *Mathematical Surveys and Monographs*, Vol. 127, [Amer. Math. Soc.](#), Providence, RI, 2006.
- [16] Cohen F.R., Lada T.J., May J.P., The homology of iterated loop spaces, *Lecture Notes in Math.*, Vol. 533, [Springer-Verlag](#), Berlin – New York, 1976.
- [17] Cornea O., Lupton G., Oprea J., Tanré D., Lusternik–Schnirelmann category, *Mathematical Surveys and Monographs*, Vol. 103, [Amer. Math. Soc.](#), Providence, RI, 2003.
- [18] Fadell E., Husseini S., Category weight and Steenrod operations, *Bol. Soc. Mat. Mexicana* **37** (1992), 151–161.
- [19] Fadell E., Husseini S., Geometry and topology of configuration spaces, *Springer Monographs in Mathematics*, [Springer-Verlag](#), Berlin, 2001.
- [20] Farber M., Topology of billiard problems. II, *Duke Math. J.* **115** (2002), 587–621, [arXiv:math.AT/0006085](#).
- [21] Farber M., Tabachnikov S., Periodic trajectories in 3-dimensional convex billiards, *Manuscripta Math.* **108** (2002), 431–437, [arXiv:math.DG/0106049](#).

- [22] Farber M., Tabachnikov S., Topology of cyclic configuration spaces and periodic trajectories of multi-dimensional billiards, *Topology* **41** (2002), 553–589, [arXiv:math.DG/9911226](#).
- [23] Fenchel W., On the differential geometry of closed space curves, *Bull. Amer. Math. Soc.* **57** (1951), 44–54.
- [24] Fomenko A., Fuchs D., Homotopical topology, 2nd ed., *Graduate Texts in Mathematics*, Vol. 273, Springer, Cham, 2016.
- [25] Franks J., Geodesics on S^2 and periodic points of annulus homeomorphisms, *Invent. Math.* **108** (1992), 403–418.
- [26] Gutkin B., Hyperbolic magnetic billiards on surfaces of constant curvature, *Comm. Math. Phys.* **217** (2001), 33–53, [arXiv:nlin.CD/0003033](#).
- [27] Gutkin E., Tabachnikov S., Billiards in Finsler and Minkowski geometries, *J. Geom. Phys.* **40** (2002), 277–301.
- [28] Hatcher A., Algebraic topology, Cambridge University Press, Cambridge, 2002.
- [29] Karasev R.N., Periodic billiard trajectories in smooth convex bodies, *Geom. Funct. Anal.* **19** (2009), 423–428, [arXiv:0905.1761](#).
- [30] Katok A.B., Ergodic perturbations of degenerate integrable Hamiltonian systems, *Izv. Akad. Nauk SSSR Ser. Mat.* **37** (1973), 539–576.
- [31] Kozlov V.V., Treshchev D.V., Billiards. A genetic introduction to the dynamics of systems with impacts, *Translations of Mathematical Monographs*, Vol. 89, Amer. Math. Soc., Providence, RI, 1991.
- [32] Laudенbach F., A Morse complex on manifolds with boundary, *Geom. Dedicata* **153** (2011), 47–57, [arXiv:1003.5077](#).
- [33] Mazzucchelli M., On the multiplicity of non-iterated periodic billiard trajectories, *Pacific J. Math.* **252** (2011), 181–205, [arXiv:1012.5593](#).
- [34] McCleary J., A user’s guide to spectral sequences, 2nd ed., *Cambridge Studies in Advanced Mathematics*, Vol. 58, Cambridge University Press, Cambridge, 2001.
- [35] Robnik M., Berry M.V., Classical billiards in magnetic fields, *J. Phys. A: Math. Gen.* **18** (1985), 1361–1378.
- [36] Roth F., On the category of Euclidean configuration spaces and associated fibrations, in Groups, Homotopy and Configuration Spaces, *Geom. Topol. Monogr.*, Vol. 13, *Geom. Topol. Publ.*, Coventry, 2008, 447–461, [arXiv:0904.1013](#).
- [37] Rudyak Yu.B., On category weight and its applications, *Topology* **38** (1999), 37–55.
- [38] Shen Y.-B., Shen Z., Introduction to modern Finsler geometry, *World Scientific Publishing Co.*, Singapore, 2016.
- [39] Strom J.A., Category weight and essential category weight, Ph.D. Thesis, The University of Wisconsin, USA, 1997.
- [40] Tabachnikov S., Billiards, *Panor. Synth.* **1** (1995), vi+142 pages.
- [41] Tabachnikov S., Remarks on magnetic flows and magnetic billiards, Finsler metrics and a magnetic analog of Hilbert’s fourth problem, in Modern Dynamical Systems and Applications, Cambridge University Press, Cambridge, 2004, 233–250, [arXiv:math.DG/0302288](#).
- [42] Tabachnikov S., Geometry and billiards, *Student Mathematical Library*, Vol. 30, Amer. Math. Soc., Providence, RI, 2005.
- [43] Tasnádi T., The behavior of nearby trajectories in magnetic billiards, *J. Math. Phys.* **37** (1996), 5577–5598.
- [44] Zharnitsky V., Invariant tori in Hamiltonian systems with impacts, *Comm. Math. Phys.* **211** (2000), 289–302.