The frequency map for billiards inside ellipsoids

Pablo S. Casas and Rafael Ramírez-Ros

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

Abstract. The billiard motion inside an ellipsoid $Q \subset \mathbb{R}^{n+1}$ is completely integrable. Its phase space is a symplectic manifold of dimension $2n$, which is mostly foliated with Liouville tori of dimension n. The motion on each Liouville torus becomes just a parallel translation with some frequency ω that varies with the torus. Besides, any billiard trajectory inside Q is tangent to n caustics $Q_{\lambda_1}, \ldots, Q_{\lambda_n}$, so the caustic parameters $\lambda = (\lambda_1, \ldots, \lambda_n)$ are integrals of the billiard map. The frequency map $\lambda \mapsto \omega$ is a key tool to understand the structure of periodic billiard trajectories. In principle, it is well-defined only for nonsingular values of the caustic parameters.

We present four conjectures, fully supported by numerical experiments. The last one gives rise to some lower bounds on the periods. These bounds only depend on the type of the caustics. We describe the geometric meaning, domain, and range of ω. The map ω can be continuously extended to singular values of the caustic parameters, although it becomes "exponentially sharp" at some of them.

Finally, we study triaxial ellipsoids of $\mathbb{R}^3$. We compute numerically the bifurcation curves in the parameter space on which the Liouville tori with a fixed frequency disappear. We determine which ellipsoids have more periodic trajectories. We check that the previous lower bounds on the periods are optimal, by displaying periodic trajectories with periods four, five, and six whose caustics have the right types. We also give some new insights for ellipses of $\mathbb{R}^2$.

Keywords: Billiards, integrability, frequency map, periodic orbits, bifurcations

AMS classification scheme numbers: 37J20, 37J35, 37J45, 70H06, 14H99

PACS numbers: 02.30.Ik, 45.20.Jj, 45.50.Tn

E-mail: pablo@casas.upc.es, Rafael.Ramirez@upc.edu

1. Introduction

Birkhoff [7] introduced the problem of *convex billiard tables* more than 75 years ago as a way to describe the motion of a free particle inside a closed convex curve such that it reflects at the boundary according to the law "angle of incidence equals angle of reflection". He also realized that this billiard motion can be modeled by an area preserving map defined on an annulus. There exists a tight relation between the invariant curves of this billiard map and the caustics of the billiard trajectories. Caustics are curves with the property that a trajectory, once tangent to it, stays tangent after every reflection. Good starting points in the literature of the billiard problem are [27, 37, 25]. We also refer to [26] for some nice figures of caustics.

When the billiard curve is an ellipse, the billiard map is integrable in the sense of Liouville, so the annulus is foliated by invariant curves, which implies that any billiard trajectory has a caustic. Indeed, the caustics of that problem are the conics confocal to the original ellipse: confocal ellipses, confocal hyperbolas, and the foci. The foci are the singular elements of the family of confocal conics. From a more dynamical point of view, the billiard map becomes just a rigid rotation on its regular invariant curves, and the rotation number varies analytically with the curve.

The billiard dynamics inside an ellipse is known. We stress just three key results related with the search of periodic trajectories. First, Poncelet showed that if a billiard trajectory is periodic, then all the trajectories sharing its caustic conic are also periodic. Second, Cayley gave some algebraic conditions to determine the caustic conics whose trajectories are periodic. Third, the rotation number can be expressed as a quotient of elliptic integrals [30, 38, 41]. We note that the search of periodic trajectories inside an ellipse can be reduced to the search of rational rotation numbers.

A rather natural generalization of this problem is to consider the motion of the particle inside an ellipsoid of $\mathbb{R}^{n+1}$. Then the phase space is no longer an annulus, but a symplectic manifold of dimension $2n$. Many of the previous results have been extended to ellipsoids, although those extensions are far from being trivial. For instance, the billiard map is still completely integrable in the sense of Liouville. Thus, the phase space is mostly foliated with Liouville tori of dimension n. The motion on each Liouville torus becomes just a parallel translation with some frequency that varies with the torus. In particular, any billiard trajectory inside an ellipsoid has n caustics, which are quadrics confocal to the original ellipsoid. This situation is fairly exceptional, since quadrics are the only smooth hypersurfaces of $\mathbb{R}^{n+1}$, $n \geq 2$, that can have caustics [6]. Some extensions of the Poncelet theorem can be found in [10, 11, 12, 33]. Several generalized Cayley-like conditions were stated in [14, 15, 16, 17]. Finally, the frequency map can be expressed in terms of hyperelliptic integrals, see [12, 32]. The setup of these last two works is $\mathbb{R}^3$, but their formulae are effortless extended to $\mathbb{R}^{n+1}$.

From Jacobi and Darboux it is known that hyperelliptic functions play a role in the description of the billiard motion inside ellipsoids and the geodesic flow on ellipsoids. Nevertheless, we skip the algebro-geometric approach (the interested reader is referred to [31, 28, 29, 2, 3]) along this paper, in order to emphasize the dynamical point of view. Here, we just mention that the billiard dynamics inside an ellipsoid can be expressed in terms of some Riemann theta-functions associated to a hyperelliptic curve, and so, one can write down explicitely the parameterizations of the invariant tori that foliate the phase space; see [39, 21].

In order to get a flavour of the kind of results obtained and problems tackled along this paper, let us consider the three-dimensional problem. Then the billiard trajectories are tangent to two quadrics of the confocal family

$$Q_\lambda = \left\{ (x,y,z) \in \mathbb{R}^3 : \frac{x^2}{a-\lambda} + \frac{y^2}{b-\lambda} + \frac{z^2}{c-\lambda} = 1 \right\}, \qquad a > b > c > 0.$$

These quadrics are ellipsoids, one-sheet hyperbolids, and two-sheet hyperboloids, although not all the combinations can take place. For instance, both caustics can not be ellipsoids. It

is known that there are just four possible combinations: E-H1, E-H2, H1-H1, and H1-H2 (the notation is self-explanatory). We raise the following questions:

- Which is the minimal period among the trajectories with caustics of a fixed type? We conjecture that the minimal period is four in the case H1-H1, five in the cases E-H1 and E-H2, and six in the case H1-H2. This conjecture is supported by numerical experiments, as well as by some theoretical arguments; see theorem 7. Samples of periodic trajectories with minimal period are displayed in figure 11.
- Which are the ellipsoids with such minimal periodic trajectories? Which are the ones with trajectories of some fixed frequency? Not all. They are the ones close to flat: $c \ll b$, in the cases E-H1 and H1-H1; the ones far from oblate: $b \ll a$, in the case E-H2; and ellipsoids with a more complicated description in the case H1-H2; see figure 10.
- Which is the asymptotic behaviour of the frequency map close to singular caustics? We shall prove that the frequency map can be continuously extended to singular values of the caustic parameters, although it becomes "exponentially sharp" at some of them; see theorems 11 and 14. This explains an amazing phenomenon that we discovered after our first numerical computations: we found that when one looks for trajectories satisfying some given property that depends just on the caustics (for instance, seven-periodic trajectories crossing exactly six times the plane $z = 0$, or quasiperiodic trajectories with a fixed Diophantine frequency vector), many times some caustic (if not both) is very close to a degenerate one. We explain the reason in remark 10.

Finally, we want to mention that there exist many remarkable results about periodic trajectories in other billiard and geodesic problems. For instance, several nice algebraic closed geodesics on a triaxial ellipsoid can be seen in [22], and a Cayley-like condition for billiards *on* quadrics was established in [1]. Some results stray from any integrable setup. For example, some general lower bounds on the number of periodic billiard trajectories inside strictly convex smooth hypersurfaces can be found in [4, 19, 20]. The planar case was already solved by Birkhoff [7]. Of course, these lower bounds are useless for integrable systems where the periodic trajectories are organized in continuous families.

We complete this introduction with a note on the organization of the paper. In section 2 we review briefly some well-known results about billiards inside ellipsoids in order to fix notations that will be used along the rest of the paper. Next, the frequency map is introduced and interpreted in section 3. This section, concerning ellipsoids of $\mathbb{R}^{n+1}$, also contains four conjectures and the lower bounds on the periods. Billiards inside ellipses of $\mathbb{R}^2$ and inside triaxial ellipsoids of $\mathbb{R}^3$ are thoroughly studied in sections 4 and 5, respectively. Billiards inside nondegenerate ellipsoids of $\mathbb{R}^{n+1}$ are revisited in section 6. Some technical lemmas have been relegated to the appendices.

2. Preliminaries

In this section details are scarce and technicalities are avoided. Experts can simply browse this section. We will list several basic references for the more novice readers.

2.1. Confocal quadrics and elliptic billiards

The following results go back to Jacobi, Chasles, Poncelet, Darboux, and Cayley.

The starting point of our discussion is the n-dimensional nondegenerate ellipsoid

$$Q = \left\{ x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} \frac{x_i^2}{a_i} = 1 \right\}, \tag{1}$$

where $a_1, \dots, a_{n+1}$ are some fixed real constants such that $0 < a_1 < \cdots < a_{n+1}$. The degenerate cases in which the ellipsoid has some symmetry of revolution are not considered here. This ellipsoid is an element of the family of *confocal quadrics* given by

$$Q_\mu = \left\{ x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} \frac{x_i^2}{a_i - \mu} = 1 \right\}, \qquad \mu \in \mathbb{R}.$$

The meaning of Q_μ is unclear in the singular cases $\mu \in \{a_1, \dots, a_{n+1}\}$. In fact, there are two natural choices for the singular confocal quadric Q_μ when $\mu = a_j$. The first choice is to define it as the n-dimensional coordinate hyperplane

$$H_j = \left\{ x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} : x_j = 0 \right\},$$

but it also makes sense to define it as the $(n-1)$-dimensional *focal quadric*

$$F_j = \left\{ x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} : x_j = 0 \text{ and } \sum_{i \neq j} \frac{x_i^2}{a_i - \mu} = 1 \right\},$$

which is contained in the hyperplane H_j. Both choices fit in the framework of elliptic billiards, but we shall use the notation $Q_{a_j} = H_j$ along this paper.

Theorem 1 ([31, 27, 2, 37])**.** *Once fixed a nondegenerate ellipsoid Q, we know that:*

(i) A generic point $x \in \mathbb{R}^{n+1}$ belongs to exactly $n+1$ distinct nonsingular confocal quadrics $Q_{\mu_0}, \dots, Q_{\mu_n}$ such that $\mu_0 < a_1 < \mu_1 < a_2 < \cdots < a_n < \mu_n < a_{n+1}$.

(ii) A generic line $\ell \subset \mathbb{R}^{n+1}$ is tangent to exactly n distinct nonsingular confocal quadrics $Q_{\lambda_1}, \dots, Q_{\lambda_n}$ such that $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, $\lambda_1 \in (-\infty, a_1) \cup (a_1, a_2)$, and $\lambda_i \in (a_{i-1}, a_i) \cup (a_i, a_{i+1})$, for $i = 2, \dots, n$.

Set $a_0 = 0$. If a generic point x is in the interior of the ellipsoid Q, then $\mu_0 > 0$, so $a_0 < \mu_0 < a_1$. In the same way, if a generic line ℓ has a transverse intersection with the ellipsoid Q, then $\lambda_1 > 0$, so $\lambda_1 \in (a_0, a_1) \cup (a_1, a_2)$. The values $\mu_0 = 0$ and $\lambda_1 = 0$ are attained just when $x \in Q$ and ℓ is tangent to Q, respectively.

We will say that $\mu = (\mu_0, \dots, \mu_n) \in \mathbb{R}^{n+1}$ are the *Jacobi elliptic coordinates* of the point $x = (x_1, \dots, x_{n+1})$. Cartesian and elliptic coordinates are linked by relations

$$x_j^2 = \frac{\prod_{i=0}^{n} (a_j - \mu_i)}{\prod_{i \neq j} (a_j - a_i)}, \qquad j = 1, \dots, n+1. \tag{2}$$

Elliptic coordinates define a coordinate system on each one of the 2^{n+1} open orthants of the space $\mathbb{R}^{n+1}$, but they become singular at the $n+1$ coordinate hyperplanes, because the map $x \mapsto \mu$ is not one-to-one in any neighborhood of these hyperplanes.

A point is generic in the sense of the theorem if and only if it is outside all coordinate hyperplanes. From the relation between Cartesian and elliptic coordinates, we deduce that when the point x tends to the hyperplane H_j, some elliptic coordinate μ_i tends to a_j. In fact, $i = j$ or $i = j - 1$, because of the inequalities $a_i < \mu_i < a_{i+1}$.

A line is generic in the sense of the theorem if and only if it is neither tangent to a singular confocal quadric‡ nor contained in a nonsingular confocal quadric.

If two lines obey the reflection law at a point $x \in Q$, then both lines are tangent to the same confocal quadrics [37]. This shows a tight relation between elliptic billiards and confocal quadrics: all lines of a billiard trajectory inside the ellipsoid Q are tangent to exactly n confocal quadrics $Q_{\lambda_1}, \dots, Q_{\lambda_n}$, which are called *caustics* of the trajectory. We will say that $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ are the *caustic parameters* of the trajectory.

Definition 1. *A billiard trajectory inside a nondegenerate ellipsoid of the Euclidean space* $\mathbb{R}^{n+1}$ *is* nonsingular *when it has* n *distinct nonsingular caustics; that is, when its caustic parameter belongs to the* nonsingular caustic space

$$\Lambda = \left\{ (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n : \begin{array}{l} 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n \\ \lambda_i \in (a_{i-1}, a_i) \cup (a_i, a_{i+1}) \textit{ for } 1 \le i \le n \end{array} \right\}. \quad (3)$$

We will only deal with nonsingular billiard trajectories along this paper. We denote the 2^n open connected components of the nonsingular caustic space as follows:

$$\Lambda_\sigma = \left\{ (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n : \begin{array}{l} 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n \\ \lambda_i \in (a_{i+\sigma_i - 1}, a_{i+\sigma_i}) \text{ for } 1 \le i \le n \end{array} \right\},$$

for $\sigma = (\sigma_1, \dots, \sigma_n) \in \{0, 1\}^n$. For instance, the first caustic Q_{λ_1} is an ellipsoid if and only if $\lambda_1 \in (a_0, a_1)$; that is, if and only if $\lambda \in \Lambda_\sigma$ with $\sigma_1 = 0$. We will draw the space Λ for ellipses and triaxial ellipsoids of $\mathbb{R}^3$ in sections 4 and 5, respectively.

Theorem 2 ([10, 11, 12, 14, 15, 33])**.** *If a nonsingular billiard trajectory is closed after* m_0 *bounces, all trajectories sharing the same caustics are also closed after* m_0 *bounces. The nonsingular billiard trajectories sharing the caustics* $Q_{\lambda_1}, \dots, Q_{\lambda_n}$ *close after* m *bounces —up to the group of symmetries* $G = (\mathbb{Z}/2\mathbb{Z})^{n+1}$ *of* Q*—, if and only if* $m \ge n + 1$ *and*

$$\operatorname{rank} \begin{pmatrix} h_{m+1} & \cdots & h_{n+2} \\ \vdots & & \vdots \\ h_{2m-1} & \cdots & h_{n+m} \end{pmatrix} < m - n, \quad (4)$$

where $\sqrt{(a_1 - s) \cdots (a_{n+1} - s)(\lambda_1 - s) \cdots (\lambda_n - s)} = h_0 + h_1 s + h_2 s^2 + \cdots$.

The group G is formed by the 2^{n+1} reflections —involutive linear transformations— with regard to coordinate subspaces. The phrase "a billiard trajectory closes after m bounces up to the group G" means that if $(q_k)_{k \in \mathbb{Z}}$ is the sequence of impact points of the trajectory, then there exists a reflection $g \in G$ such that $q_{k+m} = g(q_k)$ for all $k \in \mathbb{Z}$. Hence, billiard trajectories that close after m bounces up to the group G, close exactly after $m_0 = 2m$ bounces, because $q_{k+m_0} = q_{k+2m} = g(q_{k+m}) = g^2(q_k) = q_k$.

‡ By abuse of notation, it is said that a line is tangent to the singular confocal quadric Q_{a_j} when it is contained in the coordinate hyperplane H_j or when it passes through the focal quadric F_j.

Poncelet (first part) and Cayley (second part, although slightly modified) proved this theorem for ellipses. Darboux generalized Poncelet theorem to triaxial ellipsoids of $\mathbb{R}^3$. Later on, both results were generalized to any dimension; the *Poncelet theorem* in [10, 11, 12, 33], and the *Cayley condition* in [14, 15].

Remark 1. It can be read in the first papers on generalized Cayley-like conditions [14, 15], that billiard trajectories close after m bounces if and only if condition (4) holds. Later on, the same authors pointed out that the trajectories close "up to the central symmetry of the ellipsoid" [16]. The correct statement is the one given here§. The four-periodic billiard trajectory drawn with a continuous green line in figure 3 is a counterexample of the original statements. It is two-periodic only up to the y-axial symmetry of the ellipse, but not up to the central one. Furthermore, we shall prove in lemma 9 that it verifies the Cayley condition $h_3 = 0$, which corresponds to the values $n = 1$ and $m = 2$.

Although the Cayley-like condition (4) is purely algebraic and quite simple, it does not seem suitable for the numerical computation of caustic parameters associated to periodic trajectories with high periods, because the degree of the involved algebraic equations grows much faster than the period. From a numerical point of view, there are better approaches: the inversion of the frequency map is one of them.

2.2. Complete integrability of elliptic billiards

We recall some results obtained by Liouville, Arnold, Moser, and Knörrer.

A symplectic map $f : M \to M$ defined on a $2n$-dimensional symplectic manifold is *completely integrable* if there exist some smooth f-invariant functions $I_1, \ldots, I_n : M \to \mathbb{R}$ (the *integrals*) that are in involution —that is, whose pair-wise Poisson brackets vanish— and that are functionally independent almost everywhere on the phase space M. In this context, the map $I = (I_1, \ldots, I_n) : M \to \mathbb{R}^n$ is called the *momentum map*. A point $m \in M$ is a *regular point* of the momentum map when the n-form $\mathrm{d}I_1 \wedge \ldots \wedge \mathrm{d}I_n$ does not vanish at m. A vector $\lambda \in \mathbb{R}^n$ is a *regular value* of the momentum map when every point in the level set $I^{-1}(\lambda)$ is regular, in which case the level set is a Lagrangian submanifold of M and we say that $I^{-1}(\lambda)$ is a *regular level set*.

The following result is a discrete version of the Liouville-Arnold theorem.

Theorem 3 ([40])**.** *Any compact connected component of a regular level set $I^{-1}(\lambda)$ is diffeomorphic to $\mathbb{T}^n$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. In appropriate coordinates the restrictions of the map to this torus becomes a parallel translation $\theta \mapsto \theta + \omega$. The map $\lambda \mapsto \omega$ is smooth at the regular values of the momentum map.*

Thus the phase space M is almost foliated by Lagrangian invariant tori and the map on each torus is simply a parallel translation. These tori are called *Liouville tori*, the shift ω is the *frequency* of the torus, and the map $\lambda \mapsto \omega$ is the *frequency map*. The dynamics on a Liouville torus with frequency ω is m_0-periodic if and only if $m_0\omega \in \mathbb{Z}^n$. Liouville tori become just

§ Because all the points related by reflections $g \in G$ have the same elliptic coordinates; see (2).

invariant curves when $n = 1$, in which case the shift is usually called the *rotation number* of the invariant curve, and denoted by ρ, instead of ω.

Now, let Q be a (strictly) convex smooth surface of $\mathbb{R}^{n+1}$ diffeomorphic to the sphere $\mathbb{S}^n$, not necessarily an ellipsoid. The billiard motion inside Q can be modelled by means of a symplectic diffeomorphism defined on the $2n$-dimensional phase space

$$M = \{(q, p) \in Q \times \mathbb{S}^n : p \text{ is directed outward } Q \text{ at } q\} . \tag{5}$$

We define the billiard map $f : M \to M$, $f(q, p) = (q', p')$, as follows. The new velocity p' is the reflection of the old velocity p with respect to the tangent plane T_qQ. That is, if we decompose the old velocity as the sum of its tangent and normal components at the surface: $p = p_{\mathrm{t}} + p_{\mathrm{n}}$ with $p_{\mathrm{t}} \in T_qQ$ and $p_{\mathrm{n}} \in N_qQ$, then $p' = p_{\mathrm{t}} - p_{\mathrm{n}} = p - 2p_{\mathrm{n}}$. The new impact point q' is the intersection of the ray $\{q + \mu p' : \mu > 0\}$ with the surface Q. This intersection is unique and transverse by convexity.

Elliptic billiards fit in the frame of the Liouville-Arnold theorem.

Theorem 4 ([31, 29, 39, 2])**.** *The billiard map associated to the nondegenerate ellipsoid (1) is completely integrable and the caustic parameters $\lambda_1, \ldots, \lambda_n$ are the integrals. The set of regular values of the corresponding momentum map is given by (3).*

3. The frequency map

3.1. Definition and interpretation

The rotation number for the billiard inside an ellipse is a quotient of elliptic integrals; see [30, 10]. Explicit formulae for the frequency map of the billiard inside a triaxial ellipsoid of $\mathbb{R}^3$ can be found in [12, §III.C]. An equivalent formula is given in [32, §5]. Both formulae contain hyperelliptic integrals and they can be effortless generalized to any dimension. Since we want to avoid as many technicalities as possible, we will not talk about Riemann surfaces, basis of holomorphic differential forms, basis of homology groups, period matrices, or other objects that arise in the theory of algebraic curves. We just explain briefly why those objects appear in a natural way when dealing with elliptic billiards and we list some references.

It was clear from the original works of Jacobi and Darboux that hyperelliptic integrals play a key role in the description of geodesic flows on (and so, billiards inside) ellipsoids. For instance, the set of lines in $\mathbb{RP}^{n+1}$ that are tangent to n distinct nonsingular confocal quadrics is isomorphic to the set of real points in the Jacobian of a nonsingular hyperelliptic curve of genus n; see [31, 28, 2]. The billiard dynamics becomes linearized on this Jacobian and it is possible to obtain explicit formulae for the impact points of billiard trajectories, by solving the Abel-Jacobi inversion problem in terms of the Riemann theta-functions associated to the hyperelliptic curve. The nonsingular case $\lambda \in \Lambda$ was solved by Veselov [39]. Some singular cases were solved by Fedorov [21].

The following notations are required to define the frequency map. Once fixed the

parameters $a_1, \dots, a_{n+1}$ of the ellipsoid, and the caustic parameters $\lambda_1, \dots, \lambda_n$, we set

$$T(s) = \prod_{i=1}^{2n+1} (c_i - s), \quad \{c_1, \dots, c_{2n+1}\} = \{a_1, \dots, a_{n+1}\} \cup \{\lambda_1, \dots, \lambda_n\}.$$

If $\lambda \in \Lambda$, then $c_1, \dots, c_{2n+1}$ are pair-wise distinct and positive, so we can assume that

$$c_0 := 0 < c_1 < \cdots < c_{2n+1}. \tag{6}$$

Hence, $T(s)$ is positive in the $n+1$ open intervals (c_{2j}, c_{2j+1}), and the improper integrals

$$K_{ij} = \int_{c_{2j}}^{c_{2j+1}} \frac{s^i \mathrm{d}s}{\sqrt{T(s)}}, \qquad i = 0, \dots, n-1, \quad j = 0, \dots, n \tag{7}$$

are absolutely convergent, real, and positive. We also consider the $n+1$ column vectors

$$K_j = (K_{0j}, \dots, K_{n-1,j})^t \in \mathbb{R}^n.$$

It is known that vectors $K_1, \dots, K_n$ are linearly independent; see [24, §III.3].

Definition 2. *The* frequency map $\omega : \Lambda \to \mathbb{R}^n$ *of the billiard inside the nondegenerate ellipsoid Q is the unique solution of the system of n linear equations*

$$K_0 + 2\sum_{j=1}^{n} (-1)^j \omega_j K_j = 0. \tag{8}$$

Remark 2. Sometimes it is useful to think that the frequency ω depends on the parameter $c = (c_1, \dots, c_{2n+1}) \in \mathbb{R}^{2n+1}_+$, and not only on the caustic parameter $\lambda = (\lambda_1, \dots, \lambda_n) \in \Lambda$. In such situations, we will write $\omega = \varpi(c)$. This map $c \mapsto \varpi(c)$ is homogeneous of degree zero and analytic in the domain defined by inequalities (6). Homogeneity is deduced by performing a change of scale in the integrals (7). Hence, we can assume without loss of generality that $c_{2n+1} = a_{n+1} = 1$. Analyticity follows from the fact that the integrands in (7) are analytic with respect to the variable of integration in all the intervals of integration and with respect to c, as long as condition (6) takes place.

This definition coincides with the formulae contained in [12, 32] for $n = 2$. It is motivated by the characterization of periodic billiard trajectories contained in the next theorem. The factor 2 has been added to simplify the interpretation of the components of the frequency map, which are all positive, due to the factors $(-1)^j$.

Theorem 5 ([16, 17])**.** *The nonsingular billiard trajectories inside the nondegenerate ellipsoid Q are periodic with exactly m_j points at $Q_{c_{2j}}$ and m_j points at $Q_{c_{2j+1}}$ if and only if $m_0 K_0 + \sum_{j=1}^{n} (-1)^j m_j K_j = 0$.*

The numbers $m_0, m_1, \dots, m_n$ that appear in theorem 5 are called *winding numbers*. The nonsingular billiard trajectories with caustic parameter λ are periodic with winding numbers $m_0, m_1, \dots, m_n$ if and only if

$$\omega_j(\lambda) = \frac{m_j}{2m_0} \in \mathbb{Q}_+, \qquad j = 1, \dots, n. \tag{9}$$

We note that m_0 is the number of bounces in the ellipsoid $Q = Q_{c_0}$, so it is the period.

Remark 3. The sequence of winding numbers of a nonsingular periodic billiard trajectory contains information about how the trajectory folds in the space $\mathbb{R}^{n+1}$. The following properties can be deduced from results contained in [16, §4.1]. Here, "number of _" means "number of times that any periodic billiard trajectory with those caustic parameters do _ along one period". The intervals (c_{2j}, c_{2j+1}) with $j \neq 0$ can adopt exactly four different forms, each one giving rise to its own geometric picture.

(i) If $(c_{2j}, c_{2j+1}) = (a_j, \lambda_{j+1})$, then m_j is the number of crossings with H_j, so it is even and $m_j/2$ is the number of oscillations around the hyperplane H_j;

(ii) If $(c_{2j}, c_{2j+1}) = (\lambda_j, a_{j+1})$, then m_j is the number of crossings with H_{j+1}, so it is even and $m_j/2$ is the number of oscillations around the hyperplane H_{j+1};

(iii) If $(c_{2j}, c_{2j+1}) = (a_j, a_{j+1})$, then m_j is the number of (alternate) crossings with H_j and H_{j+1}, so it is even and $m_j/2$ is the number of rotations that the trajectory performs when projected onto the (x_j, x_{j+1})-coordinate plane π_j; and

(iv) If $(c_{2j}, c_{2j+1}) = (\lambda_j, \lambda_{j+1})$, then m_j is the number of (alternate) tangential touches with Q_{λ_j} and $Q_{\lambda_{j+1}}$, so it can be even or odd, and it is the number of oscillations between both caustics.

These four properties suggest us the following definitions, which establish the geometric meaning of the components of the frequency map. They change with the open connected components of the nonsingular caustic space.

Definition 3. *Let $\omega = (\omega_1, \ldots, \omega_n) : \Lambda \to \mathbb{R}^n$ be the frequency map.*

(i) If $(c_{2j}, c_{2j+1}) = (a_j, \lambda_{j+1})$, then $\omega_j = m_j/2m_0$ is the H_j-oscillation number*;*

(ii) If $(c_{2j}, c_{2j+1}) = (\lambda_j, a_{j+1})$, then $\omega_j = m_j/2m_0$ is the H_{j+1}-oscillation number*;*

(iii) If $(c_{2j}, c_{2j+1}) = (a_j, a_{j+1})$, then $\omega_j = m_j/2m_0$ is the π_j-rotation number*; and*

(iv) If $(c_{2j}, c_{2j+1}) = (\lambda_j, \lambda_{j+1})$, $2\omega_j = m_j/m_0$ is the $(Q_{\lambda_j}, Q_{\lambda_{j+1}})$-oscillation number.

Remark 4. It is important to notice that (only) when a m_0-periodic billiard trajectory has two caustics of the same type —that is, when some interval (c_{2j}, c_{2j+1}) falls into the fourth case—, it is possible that $m_0\omega \notin \mathbb{Z}^n$, although then $2m_0\omega \in \mathbb{Z}^n$. This is due to the factor 2 that we have added in the definition of the frequency map.

Finally, if Q_{λ_1} is not an ellipsoid —that is, if $\lambda_1 > a_1$—, then $c_1 = a_1$, and m_0 is the number of crossings with H_1, so it is even. Therefore, the following corollary holds.

Corollary 6. *Among all the nonsingular billiard trajectories inside a nondegenerate ellipsoid, only those with an ellipsoid as caustic can have odd period.*

3.2. Four conjectures

We state four conjectures, all of them supported by numerical experiments, which, if proved, would have important consequences.

Conjecture 1. *The frequency map is a local diffeomorphism; i.e., it is nondegenerate:*

$$\det\left(\frac{\partial\omega_j}{\partial\lambda_i}(\lambda)\right)_{1\leq i,j\leq n}\neq 0, \qquad \forall\lambda\in\Lambda.$$

Popov and Topalov [32] (following a methodology developed by Knörrer [29]) have shown that the frequency map is *almost everywhere* nondegenerate when Q is a triaxial ellipsoid of $\mathbb{R}^3$, although their proof deals with just two of the four connected components of Λ; namely, the components Λ_σ with $\sigma_1=0$.

A stronger version of the previous conjecture is the following one.

Conjecture 2. *The restriction of the frequency map to any open connected component of the nonsingular caustic space is a global diffeomorphism onto its image.*

We present numerical evidence for both (local and global) conjectures in the section about triaxial ellipsoids of $\mathbb{R}^3$.

The nondegeneracy of the frequency map is important because it is an essential hypothesis —although we acknowledge that it can be replaced by some weaker Rüssmann-like nondegeneracy conditions [36, §2]— in most KAM-like theorems, which are the standard tool to prove the persistence of Liouville tori under small smooth perturbations of completely integrable maps. Therefore, if conjecture 1 holds, we can conclude that most of the Liouville tori of the billiard phase space persist under small smooth perturbations of the ellipsoid.

Conjecture 3. *The components of the frequency map are ordered in a strict decreasing order and lie in the range* $(0,1/2)$*. That is,*

$$0<\omega_n(\lambda)<\cdots<\omega_1(\lambda)<1/2, \qquad \forall\lambda\in\Lambda.$$

This conjecture is supported by the numerical computation of the frequency map for thousands of random choices of $a_1,\ldots,a_{n+1},\lambda_1,\ldots,\lambda_n$ in "dimensions" $n\leq 5$. The details of the experiments for $n=2$ are presented in section 5.

Conjecture 4. *Winding numbers of nonsingular periodic billiard trajectories are ordered in a strict decreasing order. More concretely,*

$$2\leq m_n<\cdots<m_1<m_0=\text{period}.$$

Inequality $m_n\geq 2$ is immediate, because $c_{2n+1}=a_{n+1}$, so m_n is even. Inequalities $m_j\leq m_0$ for $j\geq 1$ are also immediate, because the number of crossings with any fixed hyperplane or the number of tangential touches with any fixed caustic can not exceed the number of segments of the periodic billiard trajectory. The strict inequalities $m_j<m_0$ could be also established (using the symmetries of the ellipsoid), but we skip the details, since such a small improvement is not worth the effort.

These conjectures are interrelated. First, conjecture 2 trivially implies conjecture 1. Second, conjecture 3 implies conjecture 4; see relation (9). Third, conjecture 1 jointly with conjecture 4 imply conjecture 3. To prove this, we note that conjecture 1 implies that $\omega(\Lambda)$ is open in $\mathbb{R}^n$ and $\omega^{-1}(\mathbb{Q}^n)$ is dense in Λ, whereas conjecture 4 and relation (9) imply that the strict inequalities $0<\omega_1<\cdots<\omega_n<1/2$ hold for rational frequencies. Therefore, $0\leq\omega_1(\lambda)\leq\cdots\leq\omega_n(\lambda)\leq 1/2$ for any $\lambda\in\Lambda$, but these inequalities must be strict because

$\omega(\Lambda)$ is open. Finally, we will prove in proposition 13 that when the ellipsoid is in $\mathbb{R}^3$, conjecture 1 implies conjecture 2. Probably the proof could be generalized to any dimension, but it would be rather cumbersome. It is based on a topological argument whose key point is that the sets $\omega(\Lambda_\sigma)$ are simply connected.

3.3. Lower bounds on the periods

We know from theorem 2 that the period m_0 of any nonsingular periodic billiard trajectory inside a nondegenerate ellipsoid $Q \subset \mathbb{R}^{n+1}$ verifies that $m_0 \geq n+1$. This result can be improved in several ways using the ordering of the winding numbers stated in conjecture 4. For instance, the global lower bound $m_0 \geq n+2$ follows directly. We present below more refined semi-global lower bounds, holding each one on a different open connected component of the nonsingular caustic space. They are obtained by realizing that some winding numbers must be even in agreement with the first items in remark 3. The lower bound associated to some connected component reaches the value $2n+2$, which doubles the original lower bound given in theorem 2.

Theorem 7. *Given any $\sigma = (\sigma_1, \ldots, \sigma_n) \in \{0,1\}^n$, let E_σ be the subset of $\{0, 1, \ldots, n\}$ such that: 1) $0 \in E_\sigma \Leftrightarrow \sigma_1 = 1$; 2) $j \in E_\sigma \Leftrightarrow (\sigma_j, \sigma_{j+1}) \neq (1,0)$; and 3) $n \in E_\sigma$.*

(i) If $m_0, \ldots, m_n$ are the winding numbers of a periodic trajectory with caustic parameter in Λ_σ, then m_j is even for all $j \in E_\sigma$.

(ii) If conjecture 4 holds, then the period of any periodic billiard trajectory inside a nondegenerate ellipsoid with caustic parameter in Λ_σ can not be smaller than

$$\varkappa(\sigma) := \min \left\{ m_0 : \begin{array}{l} \exists\ 2 \leq m_n < \cdots < m_0 \text{ sequence of integers} \\ \text{such that } m_j \text{ is even for any } j \in E_\sigma \end{array} \right\}.$$

(iii) Let $\mathbf{1} = (1, \ldots, 1) \in \{0,1\}^n$, $\varsigma = (\ldots, 0, 1, 0, 1, 0) \in \{0,1\}^n$, and $\sigma \in \{0,1\}^n$. Then

$$n+2 = \varkappa(\varsigma) < \varkappa(\sigma) < \varkappa(\mathbf{1}) = 2n+2, \qquad \forall\ \sigma \neq \mathbf{1}, \varsigma.$$

Proof. (i) We recall that m_j must be even in the three first cases listed in remark 3. This is the key property. For instance, m_n is always even because $c_{2n+1} = a_{n+1}$. If $\sigma_1 = 1$, then $\lambda_1 \in (a_1, a_2)$ and $c_1 = a_1$, so m_0 is even. If m_j is odd, then $(c_{2j}, c_{2j+1}) = (\lambda_j, \lambda_{j+1})$ and $\lambda_j, \lambda_{j+1} \in (a_j, a_{j+1})$, so $(\sigma_j, \sigma_{j+1}) = (1,0)$. Hence, we have seen that $(\sigma_j, \sigma_{j+1}) \neq (1,0) \Rightarrow m_j$ is even.

(ii) This follows directly from the previous item and the definition of $\varkappa(\sigma)$.

(iii) First, we note that $E_{\mathbf{1}} = \{0, \ldots, n\}$ and $E_\varsigma = \{\ldots, n-4, n-2, n\}$. Therefore, $\varkappa(\mathbf{1}) = \min\{m_0 : \exists\ 2 \leq m_n < \cdots < m_0 \text{ sequence of even numbers}\} = 2n+2$, and

$$\varkappa(\varsigma) = \min \left\{ m_0 : \begin{array}{l} \exists\ 2 \leq m_n < \cdots < m_0 \text{ sequence of} \\ \text{integers s.t. } m_n, m_{n-2}, \ldots \text{ are even} \end{array} \right\} = n+2.$$

The minimum value of m_0 among all integer sequences such that $2 \leq m_n < \cdots < m_0$ is attained at the sequence $m_j = n+2-j$, $0 \leq j \leq n$. Thus, $\varkappa(\sigma) \geq n+2$ for all $\sigma \in \{0,1\}^n$, and $\varkappa(\sigma) = n+2 \Rightarrow E_\sigma = E_\varsigma \Rightarrow \sigma = \varsigma$.

On the other hand, $E_\sigma \subset E_{\sigma'} \Rightarrow \varkappa(\sigma) \leq \varkappa(\sigma')$. Hence, $\varkappa(\sigma) \leq \varkappa(\mathbf{1}) = 2n+2$ for all $\sigma \in \{0,1\}^n$. Finally, $\varkappa(\sigma) = 2n+2 \Rightarrow E_\sigma = \{0,\dots,n\} \Rightarrow \sigma_j \neq 0 \quad \forall j \Rightarrow \sigma = \mathbf{1}$. □

Remark 5. The lower bound $\varkappa(\sigma)$ can also be explicitely computed when $\sigma \neq \mathbf{1}, \varsigma$. For instance, let $\mathbf{0} = (0,\dots,0)$ and $\bar{\varsigma} = (\dots,1,0,1,0,1)$. Then $E_{\mathbf{0}} = \{1,\dots,n\}$ and $E_{\bar{\varsigma}} = \{\dots, n-5, n-3, n-1, n\}$, so $\varkappa(\mathbf{0}) = 2n+1$ and $\varkappa(\bar{\varsigma}) = n+3$.

Remark 6. The function $\varkappa : \{0,1\}^n \to \{n+2,\dots,2n+2\}$ is surjective and has average

$$\bar{\varkappa} := 2^{-n} \sum_{\sigma \in \{0,1\}^n} \varkappa(\sigma) = 3n/2 + 2.$$

We skip the details, the proof is by induction over n. Thus, these semi-global lower bounds improve the global lower bound $n+2$ by, in average, 50%.

Now, a natural question arises. Are these semi-global lower bounds optimal? Optimal does not mean that there exists a $\varkappa(\sigma)$-periodic billiard trajectory with caustic parameter in Λ_σ inside *all* nondegenerate ellipsoids, but just inside *some* of them. And we put another question. Which are the ellipsoids with such "minimal" periodic billiard trajectories? Both questions become almost trivial for ellipses; see subsection 4.6. The case of triaxial ellipsoids of $\mathbb{R}^3$ is answered in subsection 5.5. The general case remains open, but we conjecture that all these semi-global lower bounds are optimal.

4. Billiard inside an ellipse

In this section we describe the main analytic and geometric properties of the frequency map when $n = 1$, in which case it is called rotation number and denoted by ρ. Many of these properties are old, but the observation that the the rotation number is exponentially sharp at the singular caustic parameter seems to be new. The known results can be found in the monographes [27, 37] and the papers [30, 10, 38, 41].

4.1. Confocal caustics

To simplify the exposition, we write the ellipse as

$$Q = \left\{ (x,y) \in \mathbb{R}^2 : \frac{x^2}{a} + \frac{y^2}{b} = 1 \right\}, \qquad a > b > 0,$$

where we could assume, without loss of generality, that $a = 1$; see remark 2. Then any nonsingular billiard trajectory inside Q is tangent to one confocal caustic of the form

$$Q_\lambda = \left\{ (x,y) \in \mathbb{R}^2 : \frac{x^2}{a-\lambda} + \frac{y^2}{b-\lambda} = 1 \right\},$$

where the caustic parameter λ belongs to the nonsingular caustic space‖

$$\Lambda = E \cup H, \qquad E = (0,b), \qquad H = (b,a).$$

‖ When $\lambda \to b^-$ (resp., $\lambda \to b^+$) the caustic Q_λ flattens into the region of the x-axis enclosed by (resp., outside) the foci of the ellipse Q. When $\lambda \to a^-$, the caustic flattens into the whole y-axis.

Comparing these notations with the ones used in section 2, we get that $a_1 = b$, $a_2 = a$, $\Lambda_0 = E$, and $\Lambda_1 = H$. We have chosen those names for the connected components of Λ because then Q_λ is an ellipse for $\lambda \in E$, and a hyperbola for $\lambda \in H$.

4.2. Phase portrait

We describe now the billiard dynamics inside an ellipse. This description goes back to Birkhoff [7, §VIII.12], so it is rather old and we just list the results. Concretely, we want to know how the phase space is foliated by Liouville tori (invariant curves on which the motion becomes a rigid rotation) and separatrices (invariant curves on which the motion tends to some hyperbolic periodic trajectories).

Let us put some global coordinates (φ, r) over the billiard phase space M defined in (5), just for visualization purposes. First, following Birkhoff, we parameterize the impact points on the ellipse by means of an angular coordinate $\varphi \in \mathbb{T}$. We take, for instance, $q = \gamma(\varphi) = (a^{1/2}\cos\varphi, b^{1/2}\sin\varphi)$. Second, given an outward unitary velocity $p \in \mathbb{S}$, we set $r = \langle \gamma'(\varphi), p\rangle$, and so $|r| < |\gamma'(\varphi)| = (a\sin^2\varphi + b\cos^2\varphi)^{1/2}$. Then the correspondence $(q,p) \mapsto (\varphi, r)$ allows us to identify the phase space M with the annulus

$$\mathbb{A} = \left\{(\varphi, r) \in \mathbb{T}\times\mathbb{R} : r^2 < a\sin^2\varphi + b\cos^2\varphi\right\}. \tag{10}$$

In these coordinates, the caustic parameter becomes $\lambda(\varphi, r) = (a-b)\sin^2\varphi + b - r^2$. The partition of the annulus into invariant level curves of λ is shown in figure 1.

Each regular level set contains *two* Liouville curves and represents the family of tangent lines to a fixed nonsingular caustic Q_λ. If Q_λ is an ellipse, each Liouville curve has a one-to-one projection onto the φ coordinate and corresponds to rotations around Q_λ in opposite directions, so they are invariant under f. If Q_λ is a hyperbola, then each Liouville curve corresponds to the impacts on *one* of the two pieces of the ellipse between the branches of Q_λ, so they are exchanged under f and invariant under f^2.

The singular level set $\{(\varphi, r) \in \mathbb{A} : \lambda(\varphi, r) = b\}$ gives rise to the ∞-shaped curve

$$\lambda^{-1}(b) = \left\{(\varphi, r) \in \mathbb{A} : r = \pm(a-b)^{1/2}\sin\varphi\right\},$$

which corresponds to the family of lines through the foci. This singular level set has rotation number $1/2$; see [25, page 428]. The cross points on this singular level represent the two-periodic trajectory along the major axis of the ellipse, and the eigenvalues of the differential of the billiard map at them are positive but different from one: $\mathrm{e}^{\pm h}$ with $\cosh^2 h/2 = a/b$ and $h > 0$. On the contrary, the two-periodic trajectory along the minor axis correspond to the centers of the regions inside the ∞-shaped curve, and the eigenvalues in that case are conjugate complex of modulus one: $\mathrm{e}^{\pm 2\pi\theta \mathrm{i}}$ with $\cos^2\pi\theta = b/a$ and $0 < \theta < 1/2$. Therefore, the major axis is a hyperbolic (unstable) two-periodic trajectory and the minor axis is an elliptic (stable) one. These are the only two-periodic motions. The basic results about the stability of two-periodic billiard trajectories can be found in [27, 37].

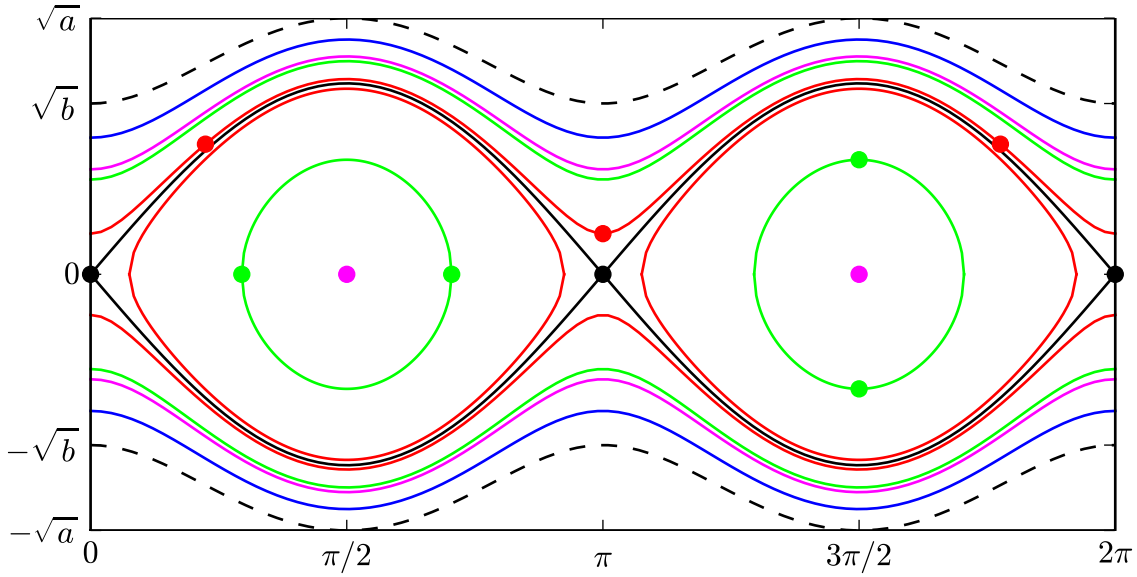

Figure 1. Phase portrait of the billiard map in (φ, r) coordinates for $a = 1$ and $b = 4/9$. The dashed black lines enclose the phase space (10). The black points are the hyperbolic two-periodic points corresponding to the oscillation along the major axis of the ellipse. The black curves are the separatrices of these hyperbolic points. The magenta points denote the elliptic two-periodic points corresponding to the oscillation along the minor axis of the ellipse. The magenta curves are the invariant curves whose rotation number coincides with the frequency of these elliptic points. The invariant curves with rotation numbers $1/6$, $1/4$ and $1/3$ are depicted in blue, green and red, respectively. The red points label a three-periodic trajectory whose caustic is an ellipse. The green points label a four-periodic trajectory whose caustic is a hyperbola.

4.3. Analytical properties of the rotation number

Let $\rho(\lambda)$ be the *rotation number* of the billiard trajectories inside the ellipse Q sharing the nonsingular caustic Q_λ. From definition 2 we get that the function $\rho : E \cup H \to \mathbb{R}$ is given by the quotients of elliptic integrals

$$\rho(\lambda) = \rho(\lambda; b, a) = \frac{\displaystyle\int_0^{\min(b,\lambda)} \frac{\mathrm{d}s}{\sqrt{(\lambda-s)(b-s)(a-s)}}}{2\displaystyle\int_{\max(b,\lambda)}^{a} \frac{\mathrm{d}s}{\sqrt{(\lambda-s)(b-s)(a-s)}}} = \frac{\displaystyle\int_\chi^\mu \frac{\mathrm{d}t}{\sqrt{t(t-1)(t-\chi)}}}{2\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{t(t-1)(t-\chi)}}}, \tag{11}$$

where the parameters $1 < \chi < \mu$ are given by $\chi = (a - \underline{m})/(a - \overline{m})$ and $\mu = a/(a - \overline{m})$, with $\underline{m} = \min(b, \lambda)$ and $\overline{m} = \max(b, \lambda)$. The second equality follows from the change of variables $t = (a - s)/(a - \overline{m})$. The second quotient already appears in [12]. Other equivalent quotients of elliptic integrals were given in [30, 41]. We have drawn the rotation function $\rho(\lambda)$ in figure 2, compare with [41, figure 2].

Proposition 8. *The function $\rho : E \cup H \to \mathbb{R}$ given in (11) has the following properties.*

(i) It is analytic in $\Lambda = E \cup H$ and increasing in E.

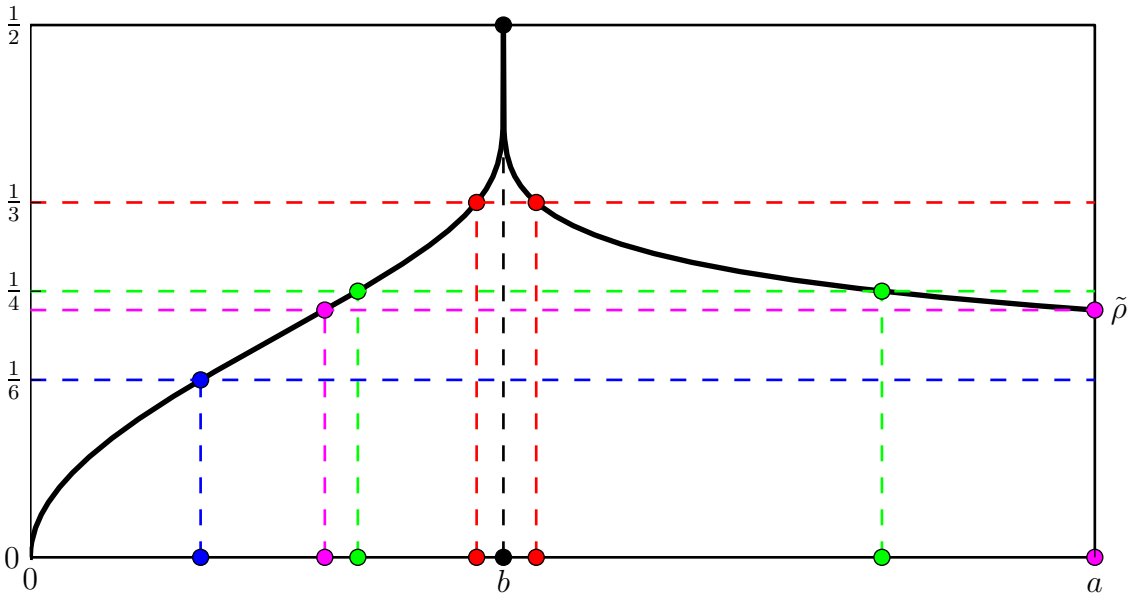

Figure 2. The rotation function $\rho(\lambda)$ of the ellipse for $a = 1$ and $b = 4/9$. Colours are taken from figure 1. The parameters $\lambda^0_\pm$ fast approach b as ρ^0 tends to $1/2$.

(ii) It can be continuously extended to the closed interval $\bar{\Lambda} = \Lambda \cup \partial\Lambda = [0, a]$ *with*

$$\rho(0) = 0, \qquad \rho(b) = 1/2, \qquad \rho(a) = \tilde{\rho},$$

where the limit value $0 < \tilde{\rho} < 1/2$ *is defined by* $\sin^2 \pi\tilde{\rho} = b/a$.

(iii) Let κ^G *and* κ^S *be the positive constants given by*

$$\kappa^G = \left(\sqrt{ab} \int_b^a \frac{\mathrm{d}s}{\sqrt{s(s-b)(a-s)}} \right)^{-1}, \qquad \cosh^2 \kappa^S = a/b.$$

The asymptotic behaviour of $\rho(\lambda)$ *at the singular parameters* $\lambda \in \partial\Lambda = \{0, b, a\}$ *is:*

(a) $\rho(\lambda) = \kappa^G \lambda^{1/2} + \mathrm{O}(\lambda^{3/2})$, *as* $\lambda \to 0^+$;

(b) $\rho(\lambda) = 1/2 + \kappa^S / \log|b - \lambda| + \mathrm{O}\left(1/\log^2|b - \lambda|\right)$, *as* $\lambda \to b$; *and*

(c) $\rho(\lambda) = \tilde{\rho} + \mathrm{O}(a - \lambda)$, *as* $\lambda \to a^-$.

(iv) Given any $\rho^0 \in (\tilde{\rho}, 1/2)$, *let* λ^0_- *be the unique parameter in* E *such that* $\rho(\lambda^0_-) = \rho^0$, *and let* λ^0_+ *be the smallest parameter in* H *such that* $\rho(\lambda^0_+) = \rho^0$. *Both parameters become exponentially close to the singular caustic parameter* b *when* ρ_0 *tends to* $1/2$. *In fact,*

$$\lambda^0_\pm = b \pm 16(a-b)\mathrm{e}^{-\kappa^S/(1/2-\rho^0)} + \mathrm{O}\left(\mathrm{e}^{-2\kappa^S/(1/2-\rho^0)}\right), \quad \rho^0 \to (1/2)^-.$$

Remark 7. We have numerically observed that the rotation function ρ is always decreasing in H, so the parameter λ^0_+ is also unique, but we have not a proof.

Proof. The analyticity follows from remark 2. The monotonicity follows from a geometric argument contained in the next subsection.

Let us write the rotation number as the quotient $\rho(\lambda) = \Delta(\lambda)/2K(\lambda)$, where

$$\Delta(\lambda) = \int_0^{\min(b,\lambda)} \frac{\mathrm{d}s}{\sqrt{T(s)}}, \qquad K(\lambda) = \int_{\max(b,\lambda)}^{a} \frac{\mathrm{d}s}{\sqrt{T(s)}},$$

and $T(s) = (\lambda - s)(b - s)(a - s)$.

The study of the limit $\lambda \to 0^+$ is easy. From lemma 17, we get the estimate $\Delta(\lambda) = 2(ab)^{-1/2}\lambda^{1/2} + \mathrm{O}(\lambda^{3/2})$, as $\lambda \to 0^+$, whereas from lemma 15 we get that

$$K(\lambda) = \int_b^a \frac{\mathrm{d}s}{\sqrt{s(s-b)(a-s)}} + \mathrm{O}(\lambda), \qquad \lambda \to 0^+.$$

By combining both estimates we get that $\rho(\lambda) = \kappa^G \lambda^{1/2} + \mathrm{O}(\lambda^{3/2})$, so $\lim_{\lambda\to 0^+} \rho(\lambda) = 0$.

Next, we consider the limit $\lambda \to b^-$. After some computations based on lemma 18, we get the estimates

$$\begin{aligned} K(b-\epsilon) &= -c^{-1/2}\log\epsilon + \eta + \mathrm{O}(\epsilon\log\epsilon), \qquad \epsilon \to 0^+, \\ \Delta(b-\epsilon) &= -c^{-1/2}\log\epsilon + \mu + \mathrm{O}(\epsilon\log\epsilon), \qquad \epsilon \to 0^+, \end{aligned}$$

where $c = a - b$, $\eta = \hat\eta + \tilde\eta$, $\mu = \hat\mu + \tilde\mu$, with $\hat\eta = c^{-1/2}\log 4c$, $\hat\mu = c^{-1/2}\log 4b$, and

$$\tilde\eta = \int_b^a \left(\frac{1}{\sqrt{a-s}} - \frac{1}{\sqrt{a-b}}\right)\frac{\mathrm{d}s}{s-b} = \frac{2}{\sqrt{c}}\int_0^{\sqrt{c}} \frac{\mathrm{d}x}{x+\sqrt{c}} = \frac{\log 4}{\sqrt{c}},$$

$$\tilde\mu = \int_0^b \left(\frac{1}{\sqrt{a-s}} - \frac{1}{\sqrt{a-b}}\right)\frac{\mathrm{d}s}{b-s} = -\frac{2}{\sqrt{c}}\int_{\sqrt{c}}^{\sqrt{a}} \frac{\mathrm{d}x}{x+\sqrt{c}} = \frac{1}{\sqrt{c}}\log\frac{4c}{(\sqrt{a}+\sqrt{c})^2}.$$

We have used the change $x^2 = a - s$ in both integrals. Let $\eta_* = c^{1/2}\eta = \log 16c$ and $\mu_* = c^{1/2}\mu = \log 16bc(a^{1/2} + c^{1/2})^{-2}$. Then we have the estimate

$$2\rho(b-\epsilon) = \frac{\Delta(b-\epsilon)}{K(b-\epsilon)} = \frac{1 - c^{1/2}\mu\log^{-1}\epsilon + \mathrm{O}(\epsilon)}{1 - c^{1/2}\eta\log^{-1}\epsilon + \mathrm{O}(\epsilon)} = \frac{1 - \mu_*\log^{-1}\epsilon}{1 - \eta_*\log^{-1}\epsilon} + \mathrm{O}(\epsilon), \tag{12}$$

as $\epsilon \to 0^+$. Thus, $\kappa^S = (\eta_* - \mu_*)/2 = \log\left((a/b)^{1/2} + (c/b)^{1/2}\right) = \log\left(d + (d^2-1)^{1/2}\right) = \operatorname{acosh} d$, where $d = (a/b)^{1/2}$. This implies that $\cosh^2\kappa^S = a/b$. Besides, estimate (12) is the key to prove that the caustic parameter λ_-^0 is exponentially close to b. Once fixed $\rho^0 \lesssim 1/2$, let $\lambda_-^0 \in E$ be the unique caustic parameter such that $\rho(\lambda_-^0) = \rho^0$, $0 < \epsilon = b - \lambda_-^0 \ll 1$, and $\delta = \log^{-1}\epsilon$. By finding $\delta^{-1} = \log\epsilon$ in estimate (12), we get

$$\log\epsilon = 1/\delta = \eta_* + \frac{\mu_* - \eta_*}{1 - 2\rho^0} + \mathrm{O}(\epsilon).$$

Using that $\kappa^S = (\eta_* - \mu_*)/2$ and $\eta_* = \log 16c$, we check that $\lambda_-^0 = b - \epsilon$, with

$$\epsilon = \mathrm{e}^{1/\delta} = \mathrm{e}^{\eta_* - 2\kappa^S/(1-2\rho^0) + \mathrm{O}(\epsilon)} = 16c\mathrm{e}^{-\kappa^S/(1/2-\rho^0)} + \mathrm{O}\left(\mathrm{e}^{-2\kappa^S/(1/2-\rho^0)}\right),$$

as $\rho^0 \to (1/2)^-$. The limit $\lambda \to b^+$ is completely analogous. We omit the computations.

With respect to the limit $\lambda \to a^-$, we note that $T(s) = (b-s)(a-s)^2 + \mathrm{O}(a-\lambda)$ uniformly in the interval $[0, b]$. Hence,

$$\Delta(\lambda) = \int_0^b \frac{\mathrm{d}s}{(a-s)\sqrt{b-s}} + \mathrm{O}(a-\lambda) = \frac{2}{\sqrt{a-b}}\operatorname{atan}\sqrt{\frac{b}{a-b}} + \mathrm{O}(a-\lambda), \quad \lambda \to a^-.$$

Besides, from lemma 16 we get the estimate $K(\lambda) = \pi(a-b)^{-1/2} + \mathrm{O}(a-\lambda)$, as $\lambda \to a^-$. Therefore, $\rho(\lambda) = \tilde{\rho} + \mathrm{O}(a-\lambda)$, as $\lambda \to a^-$, where the limit value $\tilde{\rho} \in (0, 1/2)$ is defined by $\tan^2 \pi\tilde{\rho} = b/(a-b)$. That is, $\sin^2 \pi\tilde{\rho} = b/a$. □

Remark 8. The limit rotation number $\tilde{\rho}$ is related to the conjugate complex eigenvalues $\mathrm{e}^{\pm 2\pi\theta \mathrm{i}}$ of the elliptic two-periodic orbit. Concretely, $\theta + \tilde{\rho} = 1/2$. Besides, $\tilde{\rho} = \tilde{\rho}(b/a)$ is an increasing function that tends to zero when the ellipse flattens and tends to one half when the ellipse becomes circular. That is, $\lim_{b/a \to 0^+} \tilde{\rho} = 0$, and $\lim_{b/a \to 1^-} \tilde{\rho} = 1/2$.

Definition 4. *The continuous extension* $\rho : [0, a] \to \mathbb{R}$ *is called the* (extended) rotation function *of the ellipse* Q.

4.4. Geometric properties of the rotation number

Let us assume that the billiard trajectories sharing some nonsingular caustic Q_λ are m_0-periodic, so they describe polygons with m_0 sides inscribed in the ellipse Q. Then, according to theorem 5, equation (9), and corollary 6, it turns out that $\rho(\lambda) = m_1/2m_0$ for some integers $2 \leq m_1 < m_0$ such that m_1 is always even whereas m_0 can be odd only when Q_λ is an ellipse. Besides, from the geometric interpretation of the frequency map presented in section 3, we know that: 1) If Q_λ is an ellipse, the polygons are enclosed between Q and Q_λ, and make $m_1/2$ turns around the origin; and 2) If Q_λ is a hyperbola, the polygons are contained in the region delimited by Q and the branches of Q_λ, and cross m_1 times the minor axis of the ellipse.

These interpretations can be extended to nonperiodic trajectories. Concretely,

$$\rho(\lambda) = \begin{cases} \lim\limits_{k \to +\infty} n_k/k & \text{if } \lambda \in E, \\ \frac{1}{2} \lim\limits_{k \to +\infty} l_k/k & \text{if } \lambda \in H, \end{cases}$$

where n_k (respectively, l_k) is the number of turns around the origin (respectively, crossings of the minor axis) of the first k segments of a given billiard trajectory with caustic Q_λ.

Finally, we study the monotonocity of the extended rotation function $\rho : [0, a] \to \mathbb{R}$. We would like to check that ρ is increasing in $E = (0, b)$ and decreasing in $H = (b, a)$.

We proceed by using a *reductio ad absurdum* argument. Let us assume that ρ is not increasing in E. Then there exist a couple of caustic parameters in E with the same rotation number. We know that $\rho(\lambda)$ is analytic in E and $\rho(0) \neq \rho(b)$, so it can not be constant on any subinterval of E. Hence, we can assume without loss of generality that there exist two different caustic parameters $\lambda_1, \lambda_2 \in E$ with the same *rational* rotation number. That is, there exist two different confocal ellipses Q_{λ_1} and Q_{λ_2} whose billiard trayectories have the same period and make the same number of turns around the origin per period. This is imposible because the billiard trajectories with the small ellipse as caustic rotate faster than the ones with the big ellipse. We have reduced the original question to a question about periodic trajectories to keep all arguments finite, but it was not necessary. For instance, the monotonocity of the rotation number in E could also be obtained from less geometric arguments, by applying some standard results on circle homeomorphisms and twist maps —see, for instance, [25,

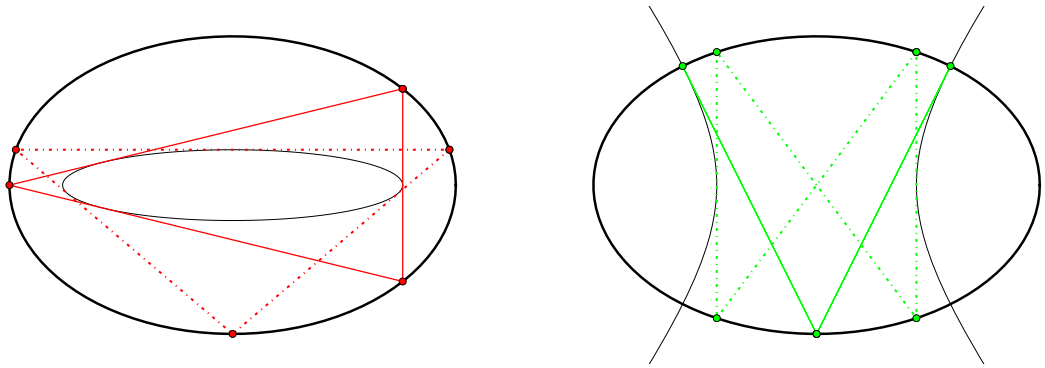

Figure 3. Examples of symmetric nonsingular billiard trajectories with minimal periods for $a = 1$ and $b = 4/9$. Left: Period three and the caustic is an ellipse. Right: Period four and the caustic is a hyperbola. The continuous lines are reserved for the trajectories that correspond to the periodic orbits depicted in figure 1.

Proposition 11.1.8]. Since the underlying idea is the same, we have preferred our geometric finite argument.

Unfortunately, neither the geometric argument nor the dynamical approach based on twist maps have a clear correspondence when the caustics are hyperbolas. The main problem is that when $\lambda \in H$ the invariant curves in the phase space (10) are not projected one-to-one onto the configuration space $\mathbb{T} \cong Q$. On the other hand, the obvious computational approach —to check that the derivative of the quotient of elliptic integrals (11) is negative for $\lambda \in H$— becomes too cumbersome. Thus, although we have no doubt that the rotation number is decreasing in H, we have failed to prove it or to find a proof in the literature.

4.5. Bifurcations in parameter space

We want to determine the ellipses that have billiard trajectories with a prescribed rotation number and with a prescribed type of caustics (ellipses or hyperbolae). We recall that the rotation function $\rho(\lambda)$ maps E onto $(0, 1/2)$, and H onto $(\tilde{\rho}, 1/2)$. Therefore, once prescribed any rotation number $\rho^0 \in (0, 1/2)$, we can always find an ellipse as caustic with that number, whereas a such hyperbola only exists when $\rho^0 > \tilde{\rho}$; that is, when $b < a \sin^2 \pi\rho^0$. This shows that flat ellipses have more periodic trajectories than rounded ones. There exists similar results for triaxial ellipsids of $\mathbb{R}^3$; see section 5.4.

4.6. Examples of periodic trajectories with minimal periods

If $n = 1$, the function $\varkappa : \{0, 1\}^n \to \{n + 2, \ldots, 2n + 2\}$ defined in theorem 7 reduces to

$$\varkappa(0) = 3, \qquad \varkappa(1) = 4.$$

We also recall that the two connected components of Λ are $\Lambda_0 = E$ and $\Lambda_1 = H$. Therefore, the lower bounds established in theorem 7 imply that the periodic trajectories with an ellipse

as caustic have period at least three, whereas the ones with a hyperbola as caustic have period at least four. In fact, those lower bounds were deduced assuming that conjecture 4 holds, but in the case $n = 1$ they can be rigorously established. It suffices to realize that the billiard map associated to an ellipse has no fixed points, its only two-periodic points correspond to the oscillations along the major or minor axis, and only the trajectories with an ellipse as caustic can have odd period.

These lower bounds are optimal; see figure 3. To be more precise, we set

$$\lambda^*_{\mathrm{E}} = \frac{3ab}{a+b+2\sqrt{a^2-ba+b^2}} \in E, \qquad \lambda^*_{\mathrm{H}} = \frac{ab}{a-b} \in H,$$

provided $b < a/2$. Then the trajectories with caustic $Q_{\lambda^*_E}$ are three-periodic, whereas the ones with caustic $Q_{\lambda^*_H}$ are four-periodic. The proof is an elementary exercise in Euclidean geometry. We leave it to the reader.

We deduce from the geometric interpretation of the rotation number given before that $\rho^*_{\mathrm{E}} = \rho(\lambda^*_{\mathrm{E}}) = 1/3$ and $\rho^*_{\mathrm{H}} = \rho(\lambda^*_{\mathrm{H}}) = 1/4$, which explains the restriction $b < a/2$; see the previous subsection. This solves the questions raised at the end of section 3.

Finally, we prove the following lemma to close the argument given in remark 1 about the problems in the original statements about generalized Cayley-like conditions.

Lemma 9. *Let* $h(s) = \sqrt{(a-s)(b-s)(\lambda-s)} = h_0 + h_1 s + h_2 s^2 + \cdots$ *for some* $0 < b < \lambda < a$, $b < a/2$. *The Cayley condition* $h_3 = 0$ *holds if and only if* $\lambda = \lambda^*_{\mathrm{H}}$.

Proof. Let $\sigma = a + b$, $p = ab$, $\sigma_1 = \lambda + \sigma$, $\sigma_2 = \sigma\lambda + p$, and $\sigma_3 = p\lambda > 0$. Then

$$(h(s))^2 = (a-s)(b-s)(\lambda-s) = -s^3 + \sigma_1 s^2 - \sigma_2 s + \sigma_3.$$

If we evaluate at $s = 0$ the first derivatives of this equality, we get that

$$h(0) = \sqrt{\sigma_3}, \qquad h'(0) = -\frac{\sigma_2}{2\sqrt{\sigma_3}}, \qquad h''(0) = \frac{4\sigma_1\sigma_3 - \sigma_2^2}{4\sigma_3\sqrt{\sigma_3}},$$

and $2h(0)h'''(0) + 6h'(0)h''(0) = -6$. Therefore,

$$\begin{aligned} h_3 = 0 &\Leftrightarrow h'''(0) = 0 \Leftrightarrow h'(0)h''(0) = -1 \Leftrightarrow 8\sigma_3^2 = \sigma_2(4\sigma_1\sigma_3 - \sigma_2^2) \\ &\Leftrightarrow (4p - \sigma^2)\sigma\lambda^3 - (4p-\sigma^2)p\lambda^2 + \sigma p^2\lambda - p^3 = 0 \\ &\Leftrightarrow (\sigma\lambda - p)\left((\sigma^2 - 4p)\lambda^2 - p^2\right) = 0 \\ &\Leftrightarrow \lambda = p/\sigma, \pm\lambda^*_{\mathrm{H}}, \end{aligned}$$

because $\sigma^2 - 4p = (a-b)^2$. These roots verify that $-\lambda^*_{\mathrm{H}} < 0 < p/\sigma < b < \lambda^*_{\mathrm{H}} < a$. □

5. Billiard inside a triaxial ellipsoid of $\mathbb{R}^3$

5.1. Confocal caustics

The caustics of a billiard inside a triaxial ellipsoid are described in several places. The representation of the caustic space shown in figure 5 can also be found in [29, 42, 18].

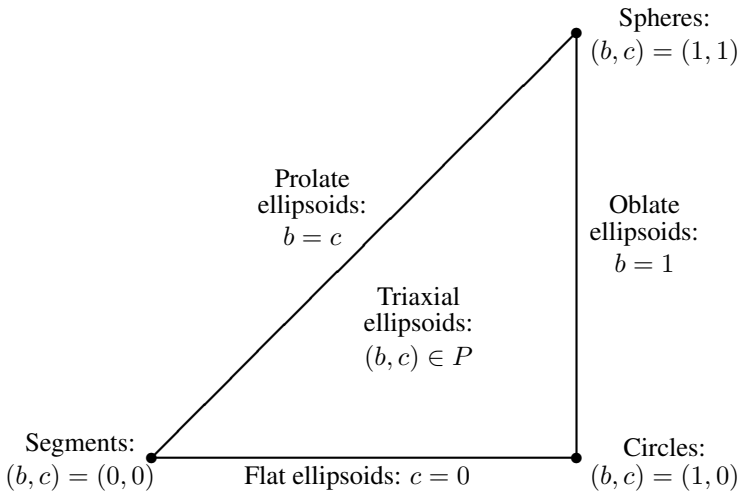

Figure 4. The triangular parameter space P.

We follow the notation for ellipses, and we write the triaxial ellipsoid as

$$Q = \left\{ (x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a} + \frac{y^2}{b} + \frac{z^2}{c} = 1 \right\}, \qquad a > b > c > 0.$$

We can assume again, without loss of generality, that $a = 1$. Then the parameter space of triaxial ellipsoids in $\mathbb{R}^3$ can be represented as the triangle

$$P = \left\{ (b, c) \in \mathbb{R}^2 : 0 < c < b < 1 \right\}, \tag{13}$$

whose edges represent ellipsoids with a symmetry of revolution (oblate and prolate ones) or flat ellipsoids, as illustrated in figure 4.

From theorem 1, we know that any nonsingular billiard trajectory inside the ellipsoid Q is tangent to *two* distinct nonsingular caustics of the confocal family

$$Q_\lambda = \left\{ (x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a - \lambda} + \frac{y^2}{b - \lambda} + \frac{z^2}{c - \lambda} = 1 \right\}.$$

The caustic Q_λ is an ellipsoid for $\lambda \in E$, a one-sheet hyperboloid when $\lambda \in H_1$, and a two-sheet hyperboloid if $\lambda \in H_2$, where

$$E = (0, c), \qquad H_1 = (c, b), \qquad H_2 = (b, a).$$

In order to have a clearer picture of how these caustics change, let us explain the situation when λ approaches the singular values c, b, or a. First, when $\lambda \to c^-$ (respectively, $\lambda \to c^+$), the caustic Q_λ flattens into the region of the coordinate plane $\pi_z = \{z = 0\}$ enclosed by (respectively, outside) the *focal ellipse*

$$Q_c^z = \left\{ (x, y, 0) \in \mathbb{R}^3 : \frac{x^2}{a - c} + \frac{y^2}{b - c} = 1 \right\}.$$

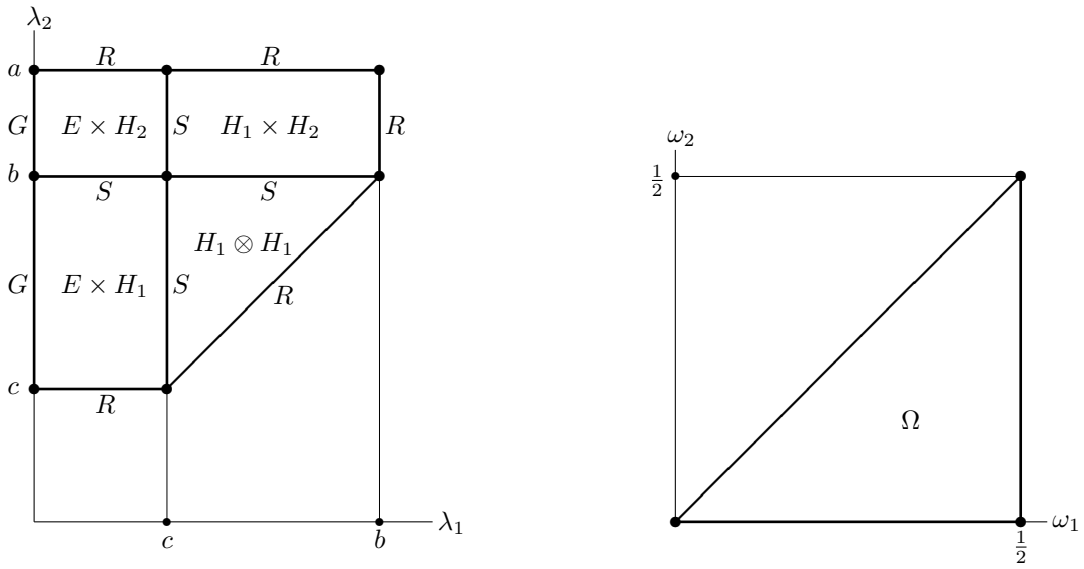

Figure 5. The nonsingular caustic space $\Lambda = (E \times H_1) \cup (E \times H_2) \cup (H_1 \otimes H_1) \cup (H_1 \times H_2)$ with its border $\partial\Lambda = G \cup R \cup S \cup \Lambda^0$ (left) and the frequency space Ω (right).

Second, when $\lambda \to b^-$ (resp., $\lambda \to b^+$), the caustic Q_λ flattens into the region of the coordinate plane $\pi_{\mathrm{y}} = \{y = 0\}$ between (resp., outside) the branches of the *focal hyperbola*

$$Q_b^{\mathrm{y}} = \left\{ (x, 0, z) \in \mathbb{R}^3 : \frac{x^2}{a-b} - \frac{z^2}{b-c} = 1 \right\}.$$

Third, the caustic flattens into the whole coordinate plane $\pi_{\mathrm{x}} = \{x = 0\}$ when $\lambda \to a^-$.

The caustic parameter $\lambda = (\lambda_1, \lambda_2)$ belongs to the nonsingular caustic space

$$\Lambda = (E \times H_1) \cup (E \times H_2) \cup (H_1 \otimes H_1) \cup (H_1 \times H_2),$$

where $H_1 \otimes H_1 = \{\lambda \in H_1 \times H_1 : \lambda_1 < \lambda_2\}$. Comparing these notations with the ones used in section 2, we get that $a_1 = c$, $a_2 = b$, $a_3 = a$, $\Lambda_{(0,0)} = E \times H_1$, $\Lambda_{(0,1)} = E \times H_2$, $\Lambda_{(1,0)} = H_1 \otimes H_1$, and $\Lambda_{(1,1)} = H_1 \times H_2$. These new notations are easier to interpret. For instance, $\lambda \in E \times H_1$ if and only if the first caustic is an ellipsoid and the second one is a one-sheet hyperboloid.

5.2. Analytical properties of the frequency map

The first analytical challenge is to determine the range of the frequency map $\omega : \Lambda \to \mathbb{R}^2$, which can be solved by studying the limits $\lim_{\lambda \to \lambda^*} \omega(\lambda)$ for $\lambda^* \in \partial\Lambda$. Nevertheless, these limits depend strongly on the "piece" of $\partial\Lambda$ under consideration. Hence, we need some notations for such "pieces".

The set Λ is the union of three open rectangles and one open isosceles rectangular triangle. In total Λ has eleven edges and eight vertexes. We consider the partitions

$$\partial\Lambda = \Lambda^0 \cup \Lambda^1, \qquad \Lambda^1 = G \cup R \cup S,$$

where Λ^1 is the set of edges, Λ^0 is the set of vertexes, and S, G and R are the sets formed by the four inner edges, the two left edges, and the remaining five edges, respectively. See the left picture of figure 5. We shall see that the frequency map is quite singular (in fact, exponentially sharp) at the four edges in S, quite regular at the five edges in R, and it is somehow related to the dynamics of the geodesic flow on the ellipsoid Q at the two edges in G. That motivates the notation.

On the other hand, if the components of the frequency map are ordered as stated in conjecture 3, the range of the frequency map must be a subset of the frequency space

$$\Omega = \left\{\omega = (\omega_1, \omega_2) \in \mathbb{R}^2 : 0 < \omega_2 < \omega_1 < 1/2\right\},$$

which is represented in the right picture of figure 5. Our main goal in this subsection is to elucidate how the the caustic space is mapped onto the frequency space. Let us summarize the final result. The frequency map folds in four the caustic space around the inner singular point $(c, b) \in \Lambda$, which is mapped onto the frequency $(1/2, 1/2)$. The vertical and horizontal "singular edges" are mapped onto the vertical and diagonal edge of the frequency space, respectively. The "geodesic edges" are mapped onto the origin.

The rest of the subsection is devoted to prove and to extend these results. First, we shall check that the frequency map of the triaxial ellipsoid Q can be continuously extended to the border of the caustic space in such a way that its values on the edges and vertexes can be expressed in terms of exactly six functions of one variable that "glue" well. Three of them are the extended rotation functions associated to the three ellipses obtained by sectioning Q with the coordinate planes π_{x}, π_{y}, and π_{z}. That is, they are the functions $\rho_{\mathrm{x}} : [0, b] \to \mathbb{R}$, $\rho_{\mathrm{y}} : [0, a] \to \mathbb{R}$, and $\rho_{\mathrm{z}} : [0, a] \to \mathbb{R}$ defined as

$$\rho_{\mathrm{x}}(\lambda) = \rho(\lambda; c, b), \qquad \rho_{\mathrm{y}}(\lambda) = \rho(\lambda; c, a), \qquad \rho_{\mathrm{z}}(\lambda) = \rho(\lambda; b, a),$$

using the notation in (11). The other three functions are defined in terms of the former ones as follows. Let $\underline{m} = \min(\lambda, c)$ and $\overline{m} = \max(\lambda, c)$. Let $T_{\mathrm{x}}(s) = (\lambda - s)(c - s)(b - s)$, $T_{\mathrm{y}}(s) = (c - s)(\lambda - s)(a - s)$, and $T_{\mathrm{z}}(s) = (\underline{m} - s)(b - s)(a - s)$. Then we consider the functions $\nu_{\mathrm{x}} : [0, b] \to \mathbb{R}$, $\nu_{\mathrm{y}} : [b, a] \to \mathbb{R}$, and $\nu_{\mathrm{z}} : [0, b] \to \mathbb{R}$ defined by the identities

$$\begin{aligned}
&\int_0^{\underline{m}} \frac{\mathrm{d}s}{(a-s)\sqrt{T_{\mathrm{x}}(s)}} - 2\rho_{\mathrm{x}}(\lambda) \int_{\overline{m}}^{b} \frac{\mathrm{d}s}{(a-s)\sqrt{T_{\mathrm{x}}(s)}} + \frac{2\pi\nu_{\mathrm{x}}(\lambda)}{\sqrt{-T_{\mathrm{x}}(a)}} = 0,\\
&\int_0^{c} \frac{\mathrm{d}s}{(b-s)\sqrt{T_{\mathrm{y}}(s)}} + 2\rho_{\mathrm{y}}(\lambda) \int_{\lambda}^{a} \frac{\mathrm{d}s}{(s-b)\sqrt{T_{\mathrm{y}}(s)}} - \frac{2\pi\nu_{\mathrm{y}}(\lambda)}{\sqrt{-T_{\mathrm{y}}(b)}} = 0,\\
&\int_0^{\underline{m}} \frac{\mathrm{d}s}{(\overline{m}-s)\sqrt{T_{\mathrm{z}}(s)}} + 2\rho_{\mathrm{z}}(\underline{m}) \int_{b}^{a} \frac{\mathrm{d}s}{(s-\overline{m})\sqrt{T_{\mathrm{z}}(s)}} - \frac{2\pi\nu_{\mathrm{z}}(\lambda)}{\sqrt{-T_{\mathrm{z}}(\overline{m})}} = 0.
\end{aligned}$$

Lemma 10. *The functions ν_{x}, ν_{y}, and ν_{z} have the following properties.*

(i) They are analytic in $E \cup H_1$, H_2, and $E \cup H_1$, respectively.

(ii) They can be continuously extended to $[0, b]$, $[b, a]$, and $[0, b]$, respectively.

(iii) Their asymptotic behaviour at the endpoints $\lambda \in \partial E \cup \partial H_1 \cup \partial H_2 = \{0, c, b, a\}$ are:

(a) $\nu_{\mathrm{x}}(\lambda) = \mathrm{O}(\lambda^{1/2})$, as $\lambda \to 0^+$;

(b) $\nu_{\mathrm{z}}(\lambda) = \mathrm{O}(\lambda^{1/2})$, *as* $\lambda \to 0^+$;
(c) $\nu_{\mathrm{x}}(\lambda) = \rho_{\mathrm{z}}(a) + \mathrm{O}\left(1/\log|c-\lambda|\right)$, *as* $\lambda \to c$;
(d) $\nu_{\mathrm{z}}(\lambda) = 1/2 + \mathrm{O}\left(|\lambda - c|^{1/2}\right)$, *as* $\lambda \to c$;
(e) $\nu_{\mathrm{x}}(\lambda) = \rho_{\mathrm{y}}(a) + \mathrm{O}(b-\lambda)$, *as* $\lambda \to b^-$;
(f) $\nu_{\mathrm{z}}(\lambda) = \rho_{\mathrm{z}}(c) + \mathrm{O}\left((b-\lambda)^{1/2}\right)$, *as* $\lambda \to b^-$;
(g) $\nu_{\mathrm{y}}(\lambda) = \rho_{\mathrm{z}}(c) + \mathrm{O}\left((\lambda-b)^{1/2}\right)$, *as* $\lambda \to b^+$; *and*
(h) $\nu_{\mathrm{y}}(\lambda) = \rho_{\mathrm{x}}(b) + \mathrm{O}(a-\lambda)$, *as* $\lambda \to a^-$.

Proof. We know that the function $\rho_{\mathrm{x}}(\lambda) = \rho(\lambda; c, b)$ is analytic in λ, c, and b, as long as $0 < c < b$ and $\lambda \in E \cup H_1$. Besides, the integrand $(a-s)^{-1}(T_{\mathrm{x}}(s))^{-1/2}$ is analytic with respect to the variable of integration s in the intervals of integration $(0, \underline{m})$ and $(\overline{m}, b)$, and with respect to the parameters λ, c, b, and a, as long as $0 < c < b < a$ and $\lambda \in E \cup H_1$. Hence, the function $\nu_{\mathrm{x}}(\lambda) = \nu_{\mathrm{x}}(\lambda; c, b, a)$ is analytic in its four variables, as long as $0 < c < b < a$ and $\lambda \in E \cup H_1$. The analyticity of ν_{y} and ν_{z} follows from similar arguments.

The study of the asymptotic behaviour of the functions ν_{x}, ν_{y}, and ν_{z} has been deferred to Appendix A.8, Appendix A.9, and Appendix A.10, respectively. □

Remark 9. We have numerically observed that ν_{x} and ν_{z} are increasing in E and decreasing in H_1, whereas ν_{y} is increasing in H_2, but we have not been able to prove it.

Theorem 11. *The frequency map* $\omega : \Lambda \to \mathbb{R}^2$ *has the following properties.*

(i) It is analytic in Λ.

(ii) It can be continuously extended to the border $\partial\Lambda$, *and the extension has the form*

$$\begin{aligned}
\omega(0, \lambda_2) &= (0, 0) && \text{for } c \le \lambda_2 \le b,\\
\omega(\lambda_1, b) &= (\rho_{\mathrm{y}}(\lambda_1), \rho_{\mathrm{y}}(\lambda_1)) && \text{for } 0 \le \lambda_1 \le b,\\
\omega(c, \lambda_2) &= (1/2, \rho_{\mathrm{z}}(\lambda_2)) && \text{for } c \le \lambda_2 \le a,\\
\omega(\lambda_1, a) &= (\rho_{\mathrm{x}}(\lambda_1), \nu_{\mathrm{x}}(\lambda_1)) && \text{for } 0 \le \lambda_1 \le b,\\
\omega(b, \lambda_2) &= (\nu_{\mathrm{y}}(\lambda_2), \rho_{\mathrm{y}}(\lambda_2)) && \text{for } b \le \lambda_2 \le a,\\
\omega(\lambda_1, c) &= (\nu_{\mathrm{z}}(\lambda_1), \rho_{\mathrm{z}}(\lambda_1)) && \text{for } 0 \le \lambda_1 \le c,\\
\omega(\lambda_1, \lambda_1) &= (\nu_{\mathrm{z}}(\lambda_1), \rho_{\mathrm{z}}(c)) && \text{for } c \le \lambda_1 \le b.
\end{aligned}$$

(iii) Its asymptotic behaviour at the eleven edges in $\Lambda^1 = G \cup S \cup R$ *is:*

(a) $\omega(\lambda_1, \lambda_2) = \kappa^G(\lambda_2)\lambda_1^{1/2} + \mathrm{O}(\lambda_1^{3/2})$, *as* $\lambda_1 \to 0^+$;
(b) $\omega(\lambda_1, \lambda_2) - \omega(c, \lambda_2) \asymp \kappa^S(c, \lambda_2)/\log|c - \lambda_1|$, *as* $\lambda_1 \to c$;
(c) $\omega(\lambda_1, \lambda_2) - \omega(\lambda_1, b) \asymp \kappa^S(\lambda_1, b)/\log|b - \lambda_2|$, *as* $\lambda_2 \to b$; *and*
(d) $\omega(\lambda) - \omega(\lambda^R) = \mathrm{O}(\lambda - \lambda^R)$, *as* $\lambda \to \lambda^R \in R$;

for some analytic functions $\kappa^G : H_1 \cup H_2 \to \mathbb{R}^2_+$ *and* $\kappa^S : S \to \mathbb{R}^2$.

(iv) Its asymptotic behaviour at the eigth vertexes in Λ^0 *is:*

(a) $\omega(\lambda_1, \lambda_2) = \mathrm{O}(\lambda_1^{1/2})$, *as* $(\lambda_1, \lambda_2) \to (0^+, c^+)$;
(b) $\omega(\lambda_1, \lambda_2) = \mathrm{O}(\lambda_1^{1/2})$, *as* $(\lambda_1, \lambda_2) \to (0^+, b)$;
(c) $\omega(\lambda_1, \lambda_2) = \mathrm{O}(\lambda_1^{1/2})$, *as* $(\lambda_1, \lambda_2) \to (0^+, a^-)$;
(d) $\omega(\lambda_1, \lambda_2) = (1/2, \rho_{\mathrm{z}}(c)) + \mathrm{O}(1/\log|c - \lambda_1|, \lambda_2 - c)$, *as* $(\lambda_1, \lambda_2) \to (c, c^+)$;

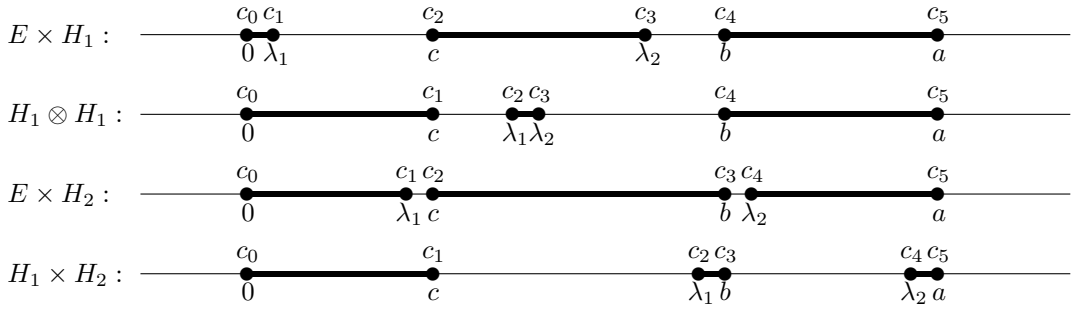

Figure 6. The four posible configurations of the ordered sequence $0 < c_1 < \cdots < c_5$. Thick lines denote intervals of integration. Each one of the displayed configurations illustrates some collapse: geodesic flow limit ($E \times H_1$), simple regular collapse ($H_1 \otimes H_1$), double singular collapse ($E \times H_2$), and double regular collapse ($H_1 \times H_2$).

(e) $\omega(\lambda_1, \lambda_2) = (1/2, 1/2) + \mathrm{O}(1/\log|c - \lambda_1|, 1/\log|b - \lambda_2|)$, *as* $(\lambda_1, \lambda_2) \to (c, b)$*;*
(f) $\omega(\lambda_1, \lambda_2) = (1/2, \rho_z(a)) + \mathrm{O}(1/\log|c - \lambda_1|, a - \lambda_2)$, *as* $(\lambda_1, \lambda_2) \to (c, a^-)$*;*
(g) $\omega(\lambda_1, \lambda_2) = (\rho_y(b), \rho_y(b)) + \mathrm{O}(b - \lambda_1, 1/\log|b - \lambda_2|)$, *as* $(\lambda_1, \lambda_2) \to (b^-, b)$*; and*
(h) $\omega(\lambda_1, \lambda_2) = (\rho_x(b), \rho_y(a)) + \mathrm{O}(b - \lambda_1, a - \lambda_2)$, *as* $(\lambda_1, \lambda_2) \to (b^-, a^-)$.

Proof. Once fixed the parameters $a > b > c > 0$ of the ellipsoid and the couple of caustic parameters λ_1 and λ_2, we set

$$\{c_1, \ldots, c_5\} = \{a, b, c\} \cup \{\lambda_1, \lambda_2\}, \qquad c_0 := 0 < c_1 < \cdots < c_5.$$

Four configurations are possible; see figure 6. We said in remark 2 that the frequency is analytic in $c_1, \ldots, c_5$ provided that $0 < c_1 < \cdots < c_5$. In particular, this implies that the frequency is analytic in the caustic parameter provided it belongs to Λ.

The frequency map is expressed in terms of six hyperelliptic integrals over the intervals $(0, c_1)$, (c_2, c_3), and (c_4, c_5) —represented in thick lines in figure 6. See definition 2. We face its asymptotic behaviour at the border $\partial\Lambda = \Lambda^0 \cup \Lambda^1$, which requires the study of the asymptotic behaviour of the six hyperelliptic integrals when some interval defined by the ordered sequence $0 < c_1 < \cdots < c_5$ collapses to a point. Therefore, there are exactly five simple collapses. The collapse of the first interval is called *geodesic flow limit*: $c_1 \to 0^+$, the collapse of the second or fourth intervals is called *singular*: $c_2 - c_1 \to 0^+$ or $c_4 - c_3 \to 0^+$, and the collapse of the third or fifth intervals is called *regular*: $c_3 - c_2 \to 0^+$ or $c_5 - c_4 \to 0^+$. Thus, regular collapses imply that the interval of integration of a couple of hyperelliptic integrals collapses to a point; whereas singular collapses imply the connection of two consecutive intervals of integration. See figure 6. It is immediate to see that this terminology agrees with the partition $\Lambda^1 = G \cup R \cup S$, whereas double collapses —that is, two simultaneous simple collapses— correspond to the eight vertexes in Λ^0.

The asymptotic behaviour of the frequency map at the eleven edges in $\Lambda^1 = G \cup R \cup S$ is deduced from several results disseminated through Appendix A. In short, some technical lemmas are listed in Appendix A.1, some notations are introduced in Appendix A.2, the geodesic flow limit is studied in Appendix A.3, simple regular collapses are analyzed

in Appendix A.4, and simple singular collapses are computed in Appendix A.5. For instance, one can trace the definition of the functions ν_{x}, ν_{y}, and ν_{z} to equation (A.3). The reader is encouraged to consult the appendix. Here, we just note that the appendix deals with the general high-dimensional setup, since the computations do not become substantially more involved when the dimension is increased.

The computations regarding the eight vertexes in Λ^0 have also been relegated to Appendix A, although it is appropiate to make a puntualization. For the sake of brevity, we have written out only the computations for two vertexes. Vertex $\lambda = (c, b)$ in Appendix A.7 —which corresponds to the unique double singular collapse—, and vertex $\lambda = (b, a)$ in Appendix A.6 —which correspond to the unique double regular collapse. The study of the remaining six vertexes does not require additional ideas. For instance, the three vertexes related to the geodesic flow limit can be simultaneously dealt with simply by using lemma 17, which ensures that the hyperelliptic integrals over $(c_0, c_1) = (0, \lambda_1)$ are $\mathrm{O}(\lambda_1^{1/2})$ as $\lambda_1 \to 0^+$.

Finally, we realize that the extended frequency map $\omega : \bar{\Lambda} \to \mathbb{R}^2$ is continuous because the extensions "glue" well at the eight vertexes; see lemma 10. For instance, let us consider the vertex (b, b). We obtain from the three statements of the theorem regarding this vertex that

$$\omega(b, b) = (\rho_{\mathrm{y}}(b), \rho_{\mathrm{y}}(b)) = (\nu_{\mathrm{y}}(b), \rho_{\mathrm{y}}(b)) = (\nu_{\mathrm{z}}(b), \rho_{\mathrm{z}}(c)),$$

which is consistent: $\nu_{\mathrm{y}}(b) = \nu_{\mathrm{z}}(b) = \rho_{\mathrm{z}}(c) = \rho(c; b, a) = \rho(b; c, a) = \rho_{\mathrm{y}}(b)$. □

Definition 5. *The continuous extension $\omega : \bar{\Lambda} \to \bar{\Omega}$ is called the* (extended) frequency map *of the ellipsoid Q.*

We are going to explain the origin of the terminology "geodesic flow limit". The phase space of the geodesic flow on an triaxial ellipsoid $Q \subset \mathbb{R}^3$ was completely described by Knörrer [28]. Any nonsingular geodesic on Q oscillates between two symmetric curvature lines obtained by intersecting Q with some hyperboloid Q_λ, $\lambda \in H_1 \cup H_2$. The rotation number of those oscillations is the quotient

$$\rho^G(\lambda) = \frac{\int_c^{\min(b,\lambda)} \frac{s\mathrm{d}s}{\sqrt{T^G(s)}}}{\int_{\max(b,\lambda)}^a \frac{s\mathrm{d}s}{\sqrt{T^G(s)}}}, \qquad T^G(s) = -s(\lambda - s)(c - s)(b - s)(a - s),$$

see [18, §4.1]. This rotation number $\rho^G(\lambda)$ can be continuously extended to the closed interval $[c, a]$ with $\rho^G(b) = 1$. On the other hand, the geodesic flow on the ellipsoid Q with caustic lines $Q \cap Q_{\lambda_2}$ can be obtained as a limit of the billiard dynamics inside Q when its first caustic Q_{λ_1} approaches Q; that is, when $\lambda_1 \to 0^+$, so that $(\lambda_1, \lambda_2) \to G$. Therefore, it is natural to look for a relation between the function $\kappa^G = (\kappa_1^G, \kappa_2^G) : H_1 \cup H_2 \to \mathbb{R}^2_+$ and the rotation number $\rho^G : H_1 \cup H_2 \to \mathbb{R}_+$.

Lemma 12. *$\rho^G = \kappa_2^G / \kappa_1^G$. Thus, $\omega_2(\lambda_1, \lambda_2)/\omega_1(\lambda_1, \lambda_2) = \rho^G(\lambda_2) + \mathrm{O}(\lambda_1)$, as $\lambda_1 \to 0^+$.*

Proof. In Appendix A.3 we will check that κ^G is the unique solution of the linear system

$$2 \begin{pmatrix} K_{01}^G & -K_{02}^G \\ K_{11}^G & -K_{12}^G \end{pmatrix} \begin{pmatrix} \kappa_1^G \\ \kappa_2^G \end{pmatrix} = \begin{pmatrix} K_{00}^G \\ 0 \end{pmatrix},$$

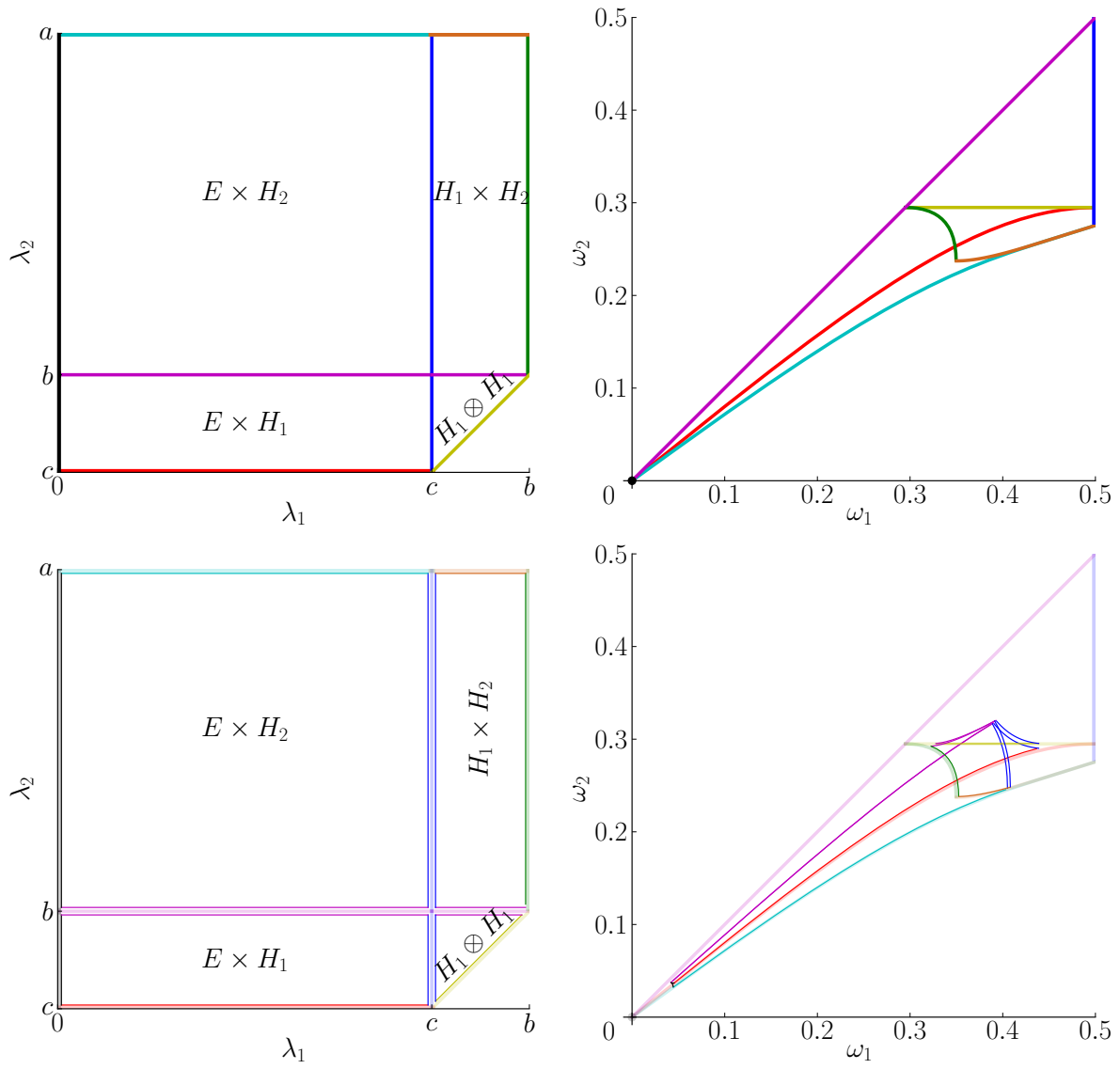

Figure 7. The extended frequency map $\omega : \bar{\Lambda} \to \bar{\Omega}$ for $a = 1$, $b = 0.58$, and $c = 0.46$. Up: On the edges of the caustic space. Down: Close to the edges of the caustic space.

where $K^G_{ij} = \int_{c_{2j}}^{c_{2j+1}} (T^G(s))^{-1/2} s^i \mathrm{d}s$, $K^G_{00} = 2(abc\lambda)^{-1/2}$, and $\{c_2, c_3, c_4, c_5\} = \{a, b, c, \lambda\}$ with $c_2 < c_3 < c_4 < c_5$. Therefore, since $\lambda \in H_1 \cup H_2$, it turns out that $c_2 = c$, $c_3 = \min(b, \lambda)$, $c_4 = \max(b, \lambda)$, and $c_5 = a$. Finally, $\kappa^G_2/\kappa^G_1 = K^G_{11}/K^G_{12} = \rho^G$. □

Next, we illustrate the results concerning the range of the frequency map and its asymptotic behaviour at the edges of the caustic space using several quantitave pictures. All the depicted curves have been numerically computed from exact formulae.

To begin with, let us consider the pictures shown in figure 7. In the upper picture we have represented the caustic space Λ at the left side, and the frequency space Ω at the right side. Each coloured segment in the caustic space is mapped onto the curve of the same colour in the frequency space. The black segment —which represents the geodesic flow limit— is mapped

onto the origin. The extended frequency map is one-to-one on each one of the red, yellow, green, brown, and cyan edges —that is, on the five regular edges. On the contrary, the magenta and blue segments are folded at their cross point (c, b). For instance, the segment joining $(0, b)$ and (c, b) is mapped onto the one joining $(0, 0)$ and $(1/2, 1/2)$, whereas the segment joining (c, b) and (b, b) is mapped onto the one joining $(1/2, 1/2)$ and $(\rho_z(c), \rho_z(c))$. The point $(\rho_z(c), \rho_z(c))$ is the intersection of the green, yellow and magenta curves in the frequency space. The triangle $H_1 \otimes H_1 \subset \Lambda$ is mapped onto the triangle $\omega(H_1 \otimes H_1) \subset \Omega$ whose edges are yellow, blue and magenta. The rectangles $E \times H_1$ and $E \times H_2$ are mapped onto regions of the frequency space with one curved and two straight borders. Finally, $\omega(H_1 \times H_2)$ has two curved and two straight borders.

In the lower picture, we have keept the upper edges and borders (drawn in light colours with thick lines), but we have added new segments and curves (drawn in heavy colours with thin lines). In the caustic space, these new segments are close to the original edges —the distance between them and the edges is equal to $c/100 = 4.6 \cdot 10^{-3}$. Nevertheless, we see that the images of the black, magenta, and blue ones are far from their corresponding borders in the frequency space. This has to do with the fact that the frequency map has an "inverse logarithm" singularity at the inner (singular) edges of the caustic space, and a "squared root" singularity at the left (geodesic) edges. This means that to reach a frequency close to the blue and magenta segments (respectively, close to the origin) one must be *exponentially close* to the inner edges (respectively, *quadratically close* to the left edges) in the caustic space.

Remark 10. This singular behaviour of the frequency map has an amazing consequence. If we look for billiard trajectories with some random frequency, it is highly probable to get some almost singular caustic. We describe a quantitative sample based on the lower picture of figure 7. If T is the triangle delimited by the yellow, blue and magenta thin segments that are close to the edges of $H_1 \otimes H_1$, the area of $\omega(H_1 \otimes H_1)$ is approximately 16 times the area of $\omega(T)$. Hence, if we look for a billiard trajectory whose caustics are one-sheet hyperboloids with a random frequency in $\omega(H_1 \otimes H_1)$, the caustic parameter $\lambda = (\lambda_1, \lambda_2)$ verifies $\min(|\lambda_1 - c|, |\lambda_2 - b|) < 4.6 \cdot 10^{-3}$ with probability 94%.

From these pictures and a topological argument whose proof has been deferred to Appendix B, we deduce that the local and global conjectures 1 and 2 are equivalent.

Proposition 13. *Conjecture 1 implies conjecture 2 and* $\omega(H_1 \otimes H_1) \subset \omega(E \times H_1)$.

Proof. Given any open connected component U of the caustic space, let $X = \partial U$ and $Y = \omega(X)$. For instance, if $U = H_1 \otimes H_1$, then X is the triangle with vertexes (c, c), (c, b), (b, b), whereas Y is the triangle with vertexes $(1/2, \rho_z(c))$, $(1/2, 1/2)$, $(\rho_z(c), \rho_z(c))$.

The borders X and Y are Jordan curves, so the frequency map $\omega : U \to \mathbb{R}^2$ verifies the hypotheses of lemma 22. Therefore, $\omega : U \to \omega(U)$ is a global diffeomorphism, being $\omega(U)$ the region enclosed inside Y.

Next, in order to prove the inclusion $\omega(H_1 \otimes H_1) \subset \omega(E \times H_1)$, it suffices to see that the red curve in the right upper picture of figure 7 is below the yellow segment. And this is

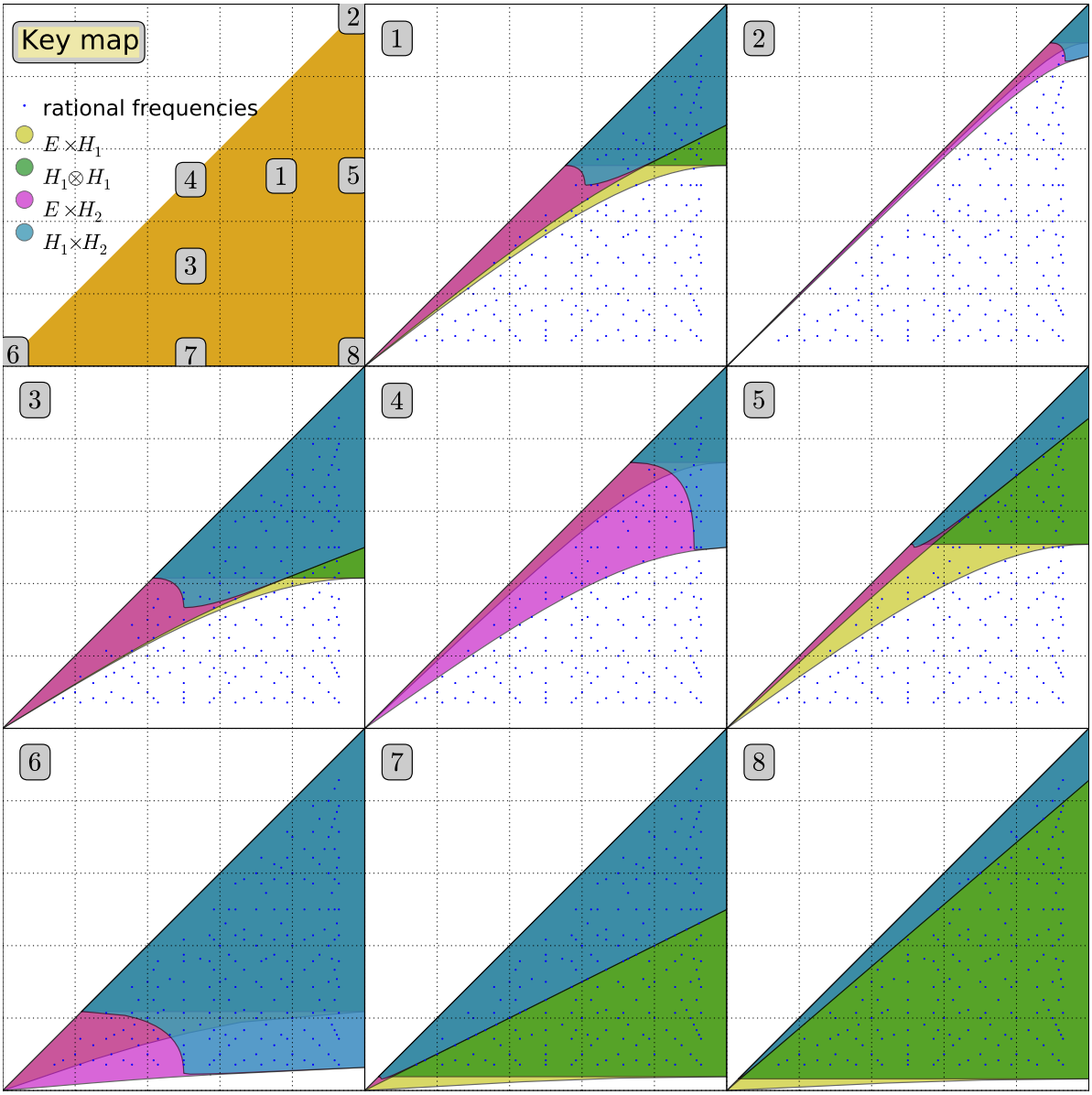

Figure 8. Ranges of the frequency map for eight different ellipsoids.

equivalent to prove the inequality

$$\rho_z(\lambda_1) \leq \rho_z(c), \qquad \forall \lambda_1 \in [0, c],$$

due to the formulae for the extended frequency map contained in theorem 11. But this inequality is a direct consequence of the monotone behaviour of the rotation number established in proposition 8. □

The ranges of the frequency map for eight different ellipsoids are shown in figure 8. In the upper left conner, we have displayed the parameter space P introduced in (13), and sketched in figure 4. We study the eight ellipsoids that correspond to the eight points in P labelled from 1) to 8). In particular, we have chosen at least one sample of each "kind" of ellipsoid: 1) standard, 2) almost spheric, 3) standard, 4) almost prolate, 5) almost oblate, 6)

close to a segment, 7) close to a flat solid ellipse, and 8) close to a flat circle. The image sets $\omega(E \times H_1)$, $\omega(H_1 \otimes H_1)$, $\omega(E \times H_2)$, and $\omega(H_1 \times H_2)$ are depicted in yellow, green, magenta, and blue, respectively. The transparency allows to visualize simultaneously all four sets. Dark blue dots correspond to rational frequencies with small common denominators. We can extract several important conclusions from these eight pictures. First, the shape of the ellipsoid has strong consequences on the range of the frequency map, which becomes extremely narrow for almost spheric ellipsoids. For instance, we see that the ellipsoids 1), 2), 4), and 5) have no periodic trajectory with frequency $\omega = (2/5, 1/5)$, which is the rational point in the frequency space Ω with the smallest common denominator. Besides, we observe that the inclusions $\omega(H_1 \otimes H_1) \subset \omega(E \times H_1)$ and $\omega(H_1 \times H_2) \subset \omega(E \times H_2)$ take place in all the computed cases. The first inclusion has already been commented in proposition 13. The second one was unexpected. To prove it, we should check that the cyan curve in the right upper picture of figure 7 is below the brown and green curves, but this requires to establish some monotonocity results about the function $\nu_{\mathrm{x}}(\lambda)$, like the one mentioned in remark 9.

We end the subsection explaining the numerical experiments about conjecture 1. We have computed the Jacobian

$$J : \Lambda \to \mathbb{R}, \qquad J(\lambda) := \det \left(\frac{\partial \omega_j}{\partial \lambda_i}(\lambda) \right)_{i,j=1,2}$$

of the frequency map for eight ellipsoids, in order to check that it never vanishes. The numerical computation —and the visualization, too— has a technical difficulty, since ω is exponentially sharp at the inner edges. To understand this fact, one can look at figure 2, where it is shown that the rotation number is exponentially sharp at $\lambda = b$. Thus, the derivative of the rotation number is exponentially big close to that point, which would make difficult its visual representation. The problem is worse in the spatial case, because the frequency map has the same kind of "inverse logarithm" singularity at the four inner edges instead of at a single point.

We overcome the visualization problem by representing the normalized Jacobian

$$J_* : \Lambda \to [0, 1], \qquad J_*(\lambda) = (1 - \exp(-|J(\lambda)|))^{1/4}.$$

The exponential function is intended to cancel the exponentially sharp behaviour of the Jacobian at the inner edges. The exponent $1/4$ has been chosen by trial and error to obtain more informative plots. The normalized Jacobian ranges over the interval $[0, 1]$. We note that $J_* = 0 \Leftrightarrow J = 0$ and $J_* = 1 \Leftrightarrow |J| = \infty$. The results are displayed in figure 9. The colour palette is a classical one: cold colours represent low values, hot colours represent high values. The neigbourhood of the inner edges is always a "hot" region; that is, the Jacobian is always (exponentially) big on that region. On the contrary, the Jacobian tends to zero close to the hypotenuse of the $H_1 \otimes H_1$ region. This can be seen from a symmetry reasoning. Furthermore, the Jacobian never vanishes, not even in the cases 7) and 8), which correspond to almost flat ellipsoids. In the left upper picture, we have again marked the ellipsoids as points in the parameter space.

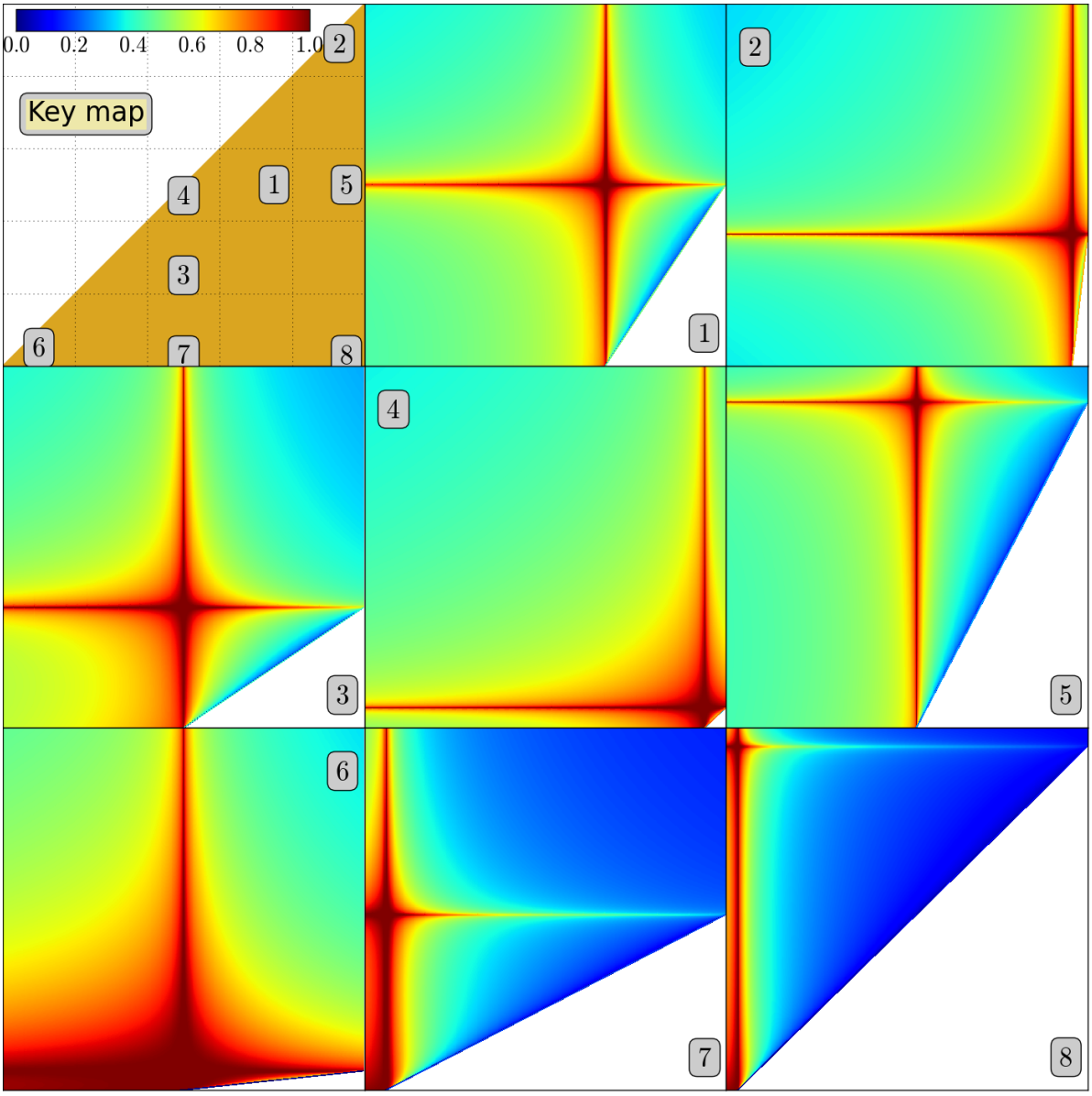

Figure 9. The normalized Jacobian $J_* : \Lambda \to [0, 1]$ of the frequency map for eight different ellipsoids.

5.3. Geometric properties of the frequency map

Let us assume that we have a periodic billiard trajectory inside the ellipsoid Q whose caustics are an ellipsoid and a one-sheet hyperboloid. Let m_0, m_1, m_2 be its winding numbers, so m_0 is its period. We deduce from remark 3 that m_1 and $m_2/2$ are the number of times along one period that the trajectory crosses the coordinate plane $\pi_z = \{z = 0\}$ and the number of times along one period that it rotates around the coordinate axis $a_z = \{x = y = 0\}$, respectively. Therefore, the components of the frequency map have the following geometric interpretation: $\omega_1 = m_1/2m_0$ is the number of oscillations around π_z per period, whereas $\omega_2 = m_2/2m_0$ is the number of rotations around a_z per period. Thus, it is quite natural to call

Type	m_1	m_2	ω_1	ω_2
$E \times H_1$	Crossings of $\pi_{\rm z}$	Half-turns around $a_{\rm z}$	z-oscillation	z-rotation
$E \times H_2$	Half-turns around $a_{\rm x}$	Crossings of $\pi_{\rm x}$	x-rotation	x-oscillation
$H_1 \otimes H_1$	Touches of Q_{λ_j}	Half-turns around $a_{\rm z}$	($\rm H_1$-oscillation)/2	z-rotation
$H_1 \times H_2$	Crossings of $\pi_{\rm y}$	Crossings of $\pi_{\rm x}$	y-oscillation	x-oscillation

Table 1. Geometric interpretation of the frequency map when $Q \subset \mathbb{R}^3$; see text.

ω_1 the z-*oscillation number* and ω_2 the z-*rotation number* of the trajectory.

As in the planar case, these interpretations are extended to nonperiodic trajectories. For instance, if $\lambda = (\lambda_1, \lambda_2) \in E \times H_1$, Q_{λ_1} is an ellipsoid, Q_{λ_2} is a one-sheet hyperboloid, and

$$\omega(\lambda) = \lim_{k \to +\infty} (n_k, l_k)/k,$$

where n_k (respectively, l_k) is the number of oscillations around $\pi_{\rm z}$ (respectively, number of rotations around $a_{\rm z}$) of the first k segments of any given trajectory with caustics Q_{λ_1} and Q_{λ_2}.

These arguments can be adapted to the other three open connected components of the caustic parameter: $E \times H_2$, $H_1 \otimes H_1$, and $H_1 \times H_2$. The results are listed in table 1. We skip the details, but a subtle point already commented in remark 4.

If the trajectory is of "type" $H_1 \otimes H_1$ —that is, if both caustics are one-sheet hyperboloids—, then the winding number m_1 is the number of (alternate) tangential touches with the caustics, so $\omega_1 = m_1/2m_0$ is half the number of oscillations between the one-sheet hyperboloids per period. We call $2\omega_1$ the $\rm H_1$-*oscillation number* of the trajectory. In particular, it can happen that $m_0\omega \notin \mathbb{Z}^2$. For instance, if the winding numbers are $m_0 = 4$, $m_1 = 3$, and $m_2 = 2$, the period is four, but $\omega = (3/8, 1/4)$.

5.4. Bifurcations in parameter space

We now extend some ideas obtained from figure 8, about the relation between the shape of the ellipsoid and the range of the frequency map. Normalizing the semimajor axis: $a = 1$, we can represent the ellipsoids by points in the parameter space P. Once fixed an open connected component Λ_σ of the caustic space and a frequency $\omega^0 = (\omega_1^0, \omega_2^0) \in \Omega$, we want to determine the ellipsoids possesing a caustic parameter $\lambda^0 \in \Lambda_\sigma$ such that $\omega(\lambda^0) = \omega^0$. Here, $\Lambda_{\mathbf{0}} = E \times H_1$, $\Lambda_{\bar\varsigma} = E \times H_2$, $\Lambda_\varsigma = H_1 \otimes H_1$, and $\Lambda_{\mathbf{1}} = H_1 \times H_2$, where $\mathbf{0} = (0,0)$, $\bar\varsigma = (0,1)$, $\varsigma = (1,0)$, and $\mathbf{1} = (1,1)$. The bifurcation curves

$$P_\sigma(\omega^0) = \big\{(b,c) \in P : \omega^0 \in \partial\omega(\Lambda_\sigma)\big\}$$

play a key role in this study. Some of them are presented in figure 10. On top of this figure we consider the eight rational frequencies with the smallest denominators. On the bottom, we depict the bifurcation curves associated to the rational frequencies marked with dark blue dots in figure 8. Next, let us describe the main properties of these curves and their endpoints. We note that although many of these properties can be analytically established, others have just been inferred from numerical evidences.

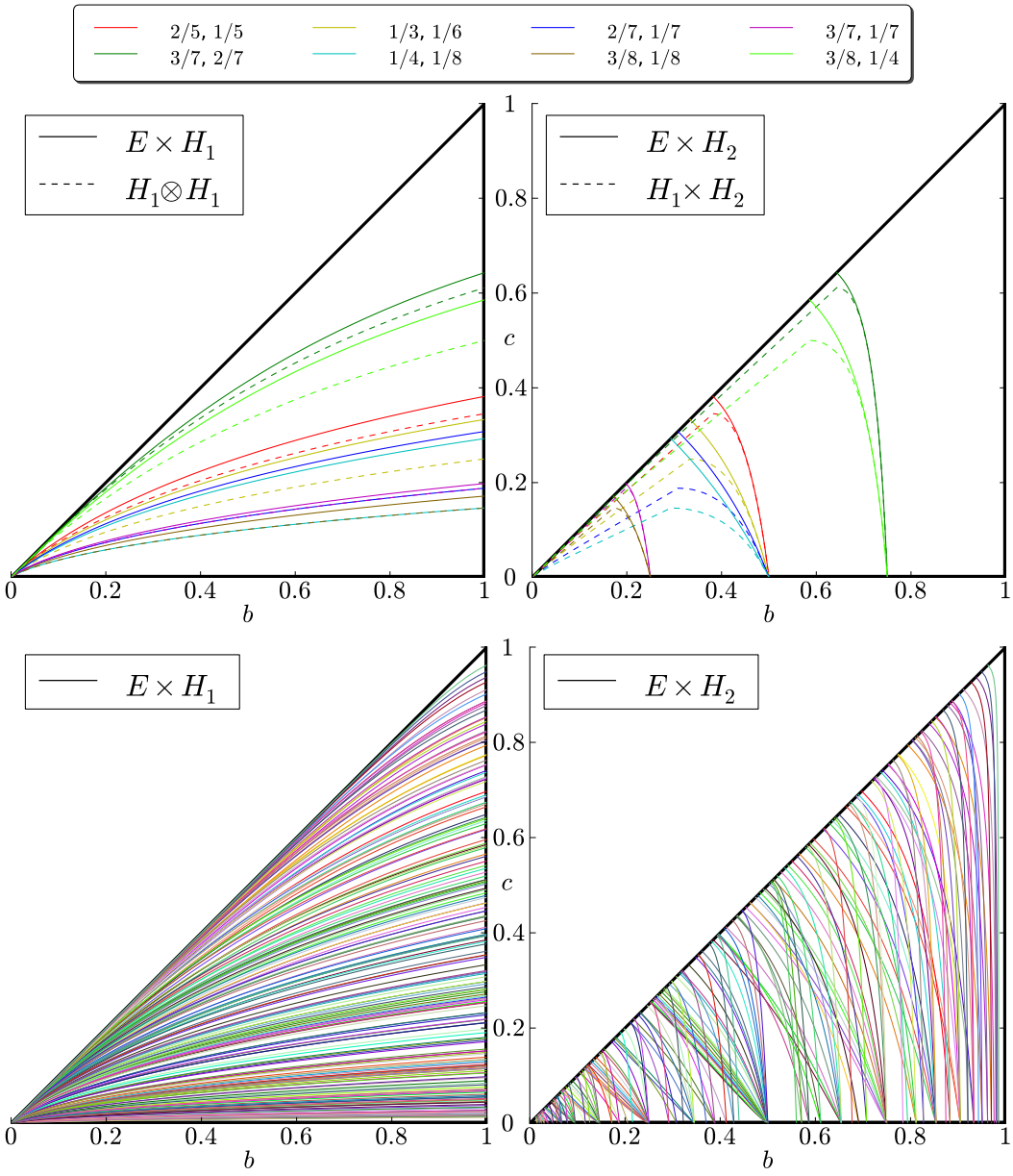

Figure 10. Some bifurcation curves $P_\sigma(\omega^0)$ in the parameter space P.

Component $\Lambda_{\mathbf{0}} = E \times H_1$. The unique moving boundary of $\omega(\Lambda_{\mathbf{0}})$ is the red curve shown in figure 7 which is parameterized by $\omega(\lambda_1, c) = (\nu_z(\lambda_1), \rho_z(\lambda_1))$, $\lambda_1 \in (0, c)$, as stated in theorem 11. Once fixed any $b \in (0, a) = (0, 1)$, we want to find λ_1 and c such that $\omega(\lambda_1, c) = \omega^0$. This looks like a root-finding problem in dimension two, but it can be written as two sequential one-dimensional root-finding problems:

(i) Find $\lambda_1 \in (0, b)$ such that $\rho_z(\lambda_1) = \rho(\lambda_1; b, a) = \omega_2^0$; and
(ii) Find $c = c_{\mathbf{0}}(b; \omega^0) \in (\lambda_1, b)$ such that $\nu_z(\lambda_1; c, b, a) = \omega_1^0$.

Then, $P_{\mathbf{0}}(\omega^0) = \{(b, c) \in P : c = c_{\mathbf{0}}(b; \omega^0)\}$. Some of these curves are depicted in the left-hand side of figure 10 with continuous lines. Each of them determine two regions

in the parameter space P; the lower one precisely describes the ellipsoids for which $\omega^0 \in \omega(E \times H_1)$. The function $c_{\mathbf{0}}(b) = c_{\mathbf{0}}(b;\omega^0)$ is increasing in $b \in (0,a)$, $c_{\mathbf{0}}(0) = 0$, and $c_{\mathbf{0}}(a) = as_2^0/s_1^0$, where $s_j^0 = \sin^2 \pi\omega_j^0$. (We recall that $a = 1$.)

Component $\Lambda_\varsigma = H_1 \otimes H_1$. This is the simplest case, because the moving boundary of $\omega(\Lambda_\varsigma)$ is the yellow horizontal segment shown in figure 7, which is parametrized by $\omega(\lambda_1,\lambda_1) = (\nu_{\mathrm{z}}(\lambda_1), \rho_{\mathrm{z}}(c))$, $\lambda_1 \in (c,b)$. Thus, in order to analyze that boundary, we just have to solve the one-dimensional equation $\rho_{\mathrm{z}}(c) = \omega_2^0$ for any fixed $b \in (0,a)$. It has a unique solution $c = c_\varsigma(b;\omega^0) \in (0,b)$, since $\rho_{\mathrm{z}}(c) = \rho(c;b,a)$ is increasing in the interval $(0,b)$, and $\rho_{\mathrm{z}}(0) = 0 < \omega_2^0 < 1/2 = \rho_{\mathrm{z}}(b)$.
Then, $P_\varsigma(\omega^0) = \{(b,c) \in P : c = c_\varsigma(b;\omega^0)\}$ does not depend on ω_1^0. This situation is observed on top left in figure 10: the two dashed curves with $\omega_2^0 = 1/8$ coincide, as well as the two ones with $\omega_2^0 = 1/7$. As in the previous case, the regions below these bifurcation curves are the ones for which $\omega^0 \in \omega(H_1 \otimes H_1)$. We observe that all dashed curves are below their corresponding continuous curves, which is in agreement with the inclusion $\omega(H_1 \otimes H_1) \subset \omega(E \times H_1)$, see proposition 13. The function $c_\varsigma(b) = c_\varsigma(b;\omega^0)$ is increasing in $b \in (0,a)$, $c_\varsigma(0) = 0$, and $c_\varsigma(a) = as_2^0$.

Component $\Lambda_{\bar\varsigma} = E \times H_2$. The moving boundary of $\omega(\Lambda_{\bar\varsigma})$ is the cyan curve shown in figure 7, parameterized by $\omega(\lambda_1,a) = (\rho_{\mathrm{x}}(\lambda_1), \nu_{\mathrm{x}}(\lambda_1))$, $\lambda_1 \in (0,c)$. Thus, the bifurcation curve $P_{\bar\varsigma}(\omega^0)$ can be computed following the ideas used in the first case. Some of these curves are depicted in the right-hand side of figure 10 with continuous lines. Regions at the left side of $P_{\bar\varsigma}(\omega^0)$ are the ones for which $\omega^0 \in \omega(E \times H_2)$. These bifurcation curves connect the prolate ellipsoid (b^*,b^*) to the flat one $(b^0,0)$, with $b^* = as_2^0/s_1^0$ and $b^0 = a\sin^2(\pi\omega_2^0/2\omega_1^0)$. To be more precise, we have numerically checked that $P_{\bar\varsigma}(\omega^0) = \{(b,c) \in P : b^* < b < b^0,\ c = c_{\bar\varsigma}(b;\omega^0)\}$ for some function $c_{\bar\varsigma}(b) = c_{\bar\varsigma}(b;\omega^0)$ decreasing in $b \in (b^*,b^0)$ such that $c_{\bar\varsigma}(b^*) = b^*$ and $c_{\bar\varsigma}(b^0) = 0$. Hence, the flat ellipsoid $(b^0,0)$ is shared by all curves having the same ratio ω_2^0/ω_1^0. This has to do with the bifurcations that take place inside an ellipse explained in subsection 4.5 through the relation $\rho^0 = \omega_2^0/2\omega_1^0$.

Component $\Lambda_{\mathbf{1}} = H_1 \times H_2$. This is the hardest case, because $\omega(\Lambda_{\mathbf{1}})$ has two moving boundaries: the brown and green curves shown in figure 7. Some of the bifurcation curves $P_{\mathbf{1}}(\omega^0)$ are depicted on top right in figure 10 with dashed lines. Each curve is composed by two arcs that glue at an interior point of P in a not differentiable way. Regions below these bifurcation curves are the ones for which $\omega^0 \in \omega(H_1 \times H_2)$. All dashed curves are below their corresponding continuous curves, which is in agreement with the inclusion $\omega(H_1 \times H_2) \subset \omega(E \times H_2)$ discussed after proposition 13.
In order to compute $P_{\mathbf{1}}(\omega^0)$, we proceed by solving separately the equations $\omega(b,\lambda_2) = \omega^0$ for $\lambda_2 \in (b,a)$, and $\omega(\lambda_1,a) = \omega^0$ for $\lambda_1 \in (c,b)$. Each equation corresponds to a moving boundary and provides an arc of $P_{\mathbf{1}}(\omega^0)$. The two arcs join at the point $(b^*,c^*) \in P$ where both equations hold simultaneously: $b^* = as_2^0/s_1^0$ and $c^* = as_2^0$, because $\omega(b,a) = (\rho_{\mathrm{x}}(b), \rho_{\mathrm{y}}(a))$, $\sin^2 \pi\rho_{\mathrm{x}}(b) = c/b$, and $\sin^2 \pi\rho_{\mathrm{y}}(a) = c/a$. More precisely, $P_{\mathbf{1}}(\omega^0) = \{(b,c) \in P : 0 < b < b^0,\ c = c_{\mathbf{1}}(b;\omega^0)\}$ for some function $c_{\mathbf{1}}(b) =$

$c_{\mathbf{1}}(b;\omega^0)$ increasing in $(0,b^*)$ and decreasing in (b^*,b^0) such that $c_{\mathbf{1}}(0)=0$, $c_{\mathbf{1}}(b^*)=c^*$, and $c_{\mathbf{1}}(b^0)=0$. Thus, $P_{\mathbf{1}}(\omega^0)$ begins at the degenerate ellipsoid $(0,0)$, passes through the triaxial ellipsoid (b^*,c^*), and ends at the flat ellipsoid $(b^0,0)$. The flat ellipsoid is again shared by all bifurcation curves having the same ratio ω_2^0/ω_1^0, see figure 10.

These properties of the bifurcation curves $P_\sigma(\omega^0)$ allow us to determine the shape of the ellipsoids with prescribed frequency ω^0 and prescribed caustic type σ. These shapes were described in the introduction.

The asymptotic properties of the functions $c_\sigma(b)$ can be rigurously established by passing to the limits the implicit equations that determine them. As a sample, the proof for the function $c_{\mathbf{0}}(b)$ can be found in Appendix A.11, the other three cases are similar. On the contrary, we do not know how to establish analytically their monotone character. Such monotone characters have always been hard challenges along this paper.

Remark 11. We have needed a multiple precision aritmethic to compute the birfucation curves close to some of their endpoints, since the involved root-finding problems become quite singular at them. The programs have been written using the PARI system [5].

5.5. Examples of periodic trajectories with minimal periods

In the case $n=2$, the function $\varkappa:\{0,1\}^n\to\{n+2,\dots,2n+2\}$ defined in theorem 7 can be computed quite easily. In fact,

$$4=\varkappa(\varsigma)<\varkappa(\mathbf{0})=5=\varkappa(\bar\varsigma)<\varkappa(\mathbf{1})=6$$

where $\varsigma=(1,0)$, $\mathbf{0}=(0,0)$, $\bar\varsigma=(0,1)$, and $\mathbf{1}=(1,1)$. Therefore, the lower bounds established in theorem 7 imply that the periodic trajectories with an ellipsoid and a hyperboloid as caustics have period at least five, the ones with two different one-sheet hyperboloids as caustics have period at least four, and the ones with two hyperboloids of different type as caustics have period at least six. We have numerically computed some symmetric periodic trajectories to check that these lower bounds are optimal; see figure 11. We shall classify all the symmetric periodic billiard trajectories inside a triaxial ellipsoid and explain how to compute them in a future paper [9].

Considering the values given in figure 11, and bearing in mind table 1, we have $(m_1,m_2)=(4,2)$ for the $E\times H_2$ trajectory, so it performs two turns around the coordinate axis a_{x} and crosses twice the coordinate plane π_{x}. As well, $(m_1,m_2)=(4,2)$ for the $E\times H_1$ trajectory, meaning four crossings with π_{z} and just one turn around a_{z}. Again, for $H_1\times H_2$ we have $(m_1,m_2)=(4,2)$, meaning four crossings with π_{y} and two crossings with π_{x}. Finally, $(m_1,m_2)=(3,2)$ for $H_1\otimes H_1$, which corresponds to three tangential touches with each of the caustics and a single turn around a_{z}. Each of those geometric interpretations has been verified on the corresponding trajectory.

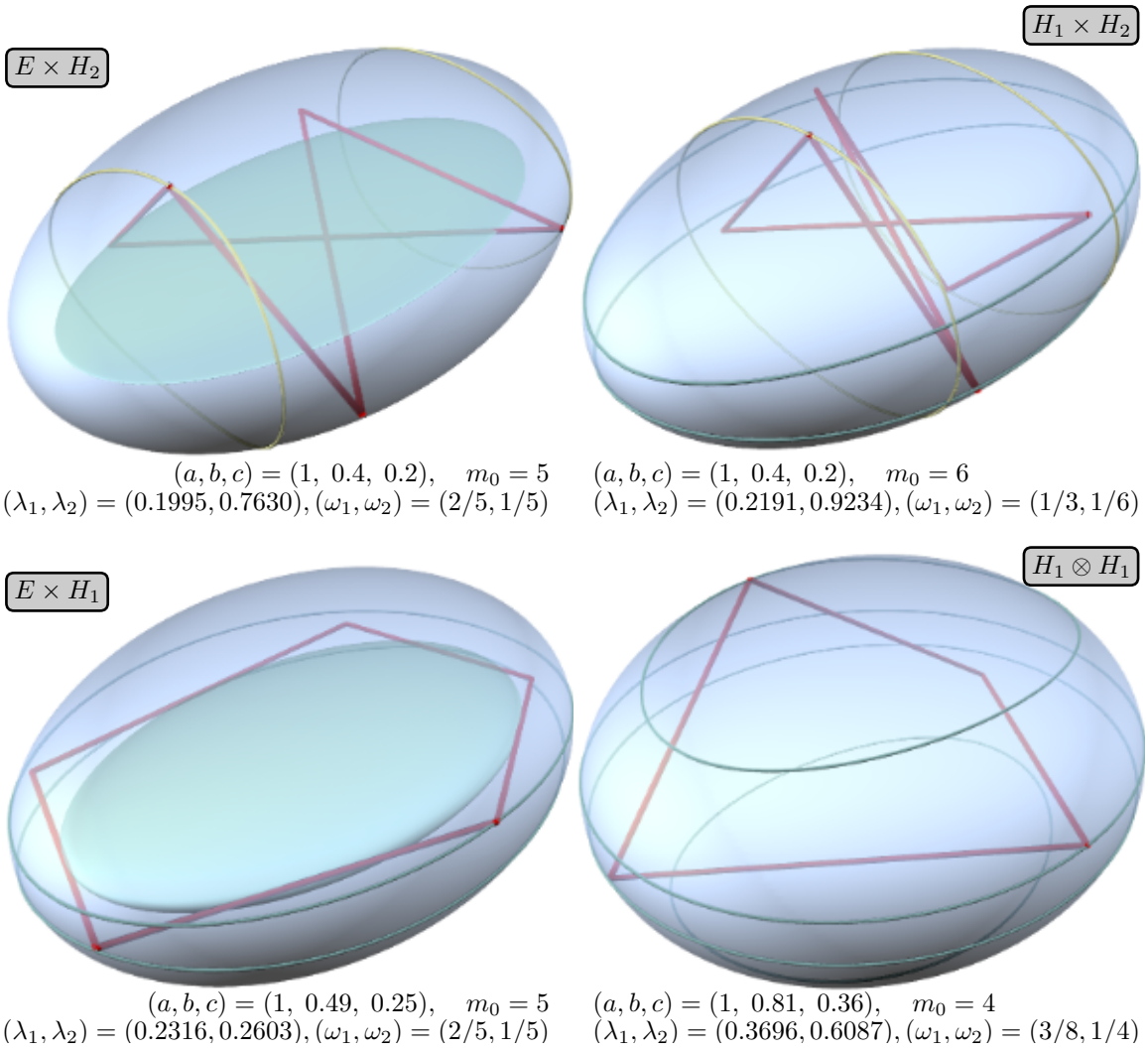

Figure 11. Examples of symmetric nonsingular billiard trajectories with minimal periods m_0. Lines in red represent the particle's trajectory. Lines in green correspond to $Q \cap H_1$, and the ones in yellow to $Q \cap H_2$. In the cases $E \times H_*$, the caustic elipsoid is also depicted.

6. Billiard inside a nondegenerate ellipsoid of $\mathbb{R}^{n+1}$

We describe briefly the high-dimensional version of some of the analytical results already shown in the spatial case. We denote again the nondegenerate ellipsoid as in (1) and the nonsingular caustic space as in (3).

By analogy with the spatial case, we consider three disjoint partitions:

$$\partial\Lambda = \cup_{k=0}^{n-1}\Lambda^k, \qquad \Lambda^{n-1} = G \cup R \cup S, \qquad S = \cup_{j=1}^{n} S_j.$$

With regard to the first one, Λ^k is the k-dimensional border of Λ. That is, Λ^0 is the set of vertexes, Λ^1 is the set of edges, Λ^2 is the set of faces, and so on. The second one mimics the distinction among geodesic flow limits, simple regular collapses, and simple singular collapses already seen in the previous section. For instance, $G = \{\lambda \in \Lambda^{n-1} : \lambda_1 = 0\}$. The

asymptotic behaviour of the frequency map in each one of these three situations is expected to be dramatically different; see the theorem below. The last partition labels the component of the caustic parameter that becomes singular: $S_j = \{\lambda \in \Lambda^{n-1} : \lambda_j = a_j\}$. Besides, given any caustic parameter $\lambda \in \Lambda$ we shall denote by $\lambda^{S_j} \in S_j$ the caustic parameter obtained from λ by substituying its j-th component with a_j. Finally, we introduce the $(n-1)$-dimensional set

$$G_* = \{(\lambda_2, \dots, \lambda_n) \in \mathbb{R}^{n-1} : (0, \lambda_2, \dots, \lambda_n) \in G\},$$

which turns out to be the nonsingular caustic space for the geodesic flow on the ellipsoid. We note that $S_1 = \{c\} \times (H_1 \cup H_2)$, $S_2 = (E \cup H_1) \times \{b\}$, and $G_* = H_1 \cup H_2$ with the notations used in the previous section for triaxial ellipsoids of $\mathbb{R}^3$.

Theorem 14. *The frequency map $\omega : \Lambda \to \mathbb{R}^n$ has the following properties.*

(i) It is analytic in Λ.

(ii) It can be continuously extended to the border $\partial\Lambda$, the extended map being as follows:

- *(a) It vanishes at $\bar{G}$;*
- *(b) One of its components can be explicitely written as a function of the rest at $\bar{R}$;*
- *(c) Its first component is equal to $1/2$ at $\bar{S}_1$;*
- *(d) Its l-th component is equal to the $(l-1)$-th component at $\bar{S}_l$ for $2 \le l \le n$; and*
- *(e) Its "free" components are an $(n-1)$-dimensional frequency of the billiard inside the section of the original ellipsoid by a suitable coordinate hyperplane at $\bar{R} \cup \bar{S}$.*

Besides, the restriction of the continuous extended map to any of the k-dimensional connected components of Λ^k, $1 \le k \le n-1$, is analytic.

(iii) Its asymptotic behaviour at $\Lambda^{n-1} = G \cup S \cup R$ is:

- *(a) $\omega(\lambda) = \kappa^G(\lambda_2, \dots, \lambda_n)\lambda_1^{1/2} + \mathrm{O}(\lambda_1^{3/2})$, as $\lambda_1 \to 0^+$;*
- *(b) $\omega(\lambda) - \omega(\lambda^{S_j}) \asymp \kappa^S(\lambda^{S_j})/\log|a_j - \lambda_j|$, as $\lambda_j \to a_j$;*
- *(c) $\omega(\lambda) - \omega(\lambda^R) = \mathrm{O}(\lambda - \lambda^R)$, as $\lambda \to \lambda^R \in R$;*

for some analytic functions $\kappa^G : G_* \to \mathbb{R}^n_+$ and $\kappa^S : S \to \mathbb{R}^n$.

Proof. It follows from the same arguments and computations that in the spatial case. The arguments are not repeated. The computations with hyperelliptic integrals have been relegated to Appendix A. □

We recall that, once fixed the parameters $a_1, \dots, a_{n+1}$ of the ellipsoid and the caustic parameters $\lambda_1, \dots, \lambda_n$, we write the $2n+1$ positive numbers

$$\{c_1, \dots, c_{2n+1}\} = \{a_1, \dots, a_{n+1}\} \cup \{\lambda_1, \dots, \lambda_n\}$$

in an ordered way: $c_0 := 0 < c_1 < \cdots < c_{2n+1}$. Then the frequency $\omega(\lambda)$ is defined in terms of some hyperelliptic integrals over the intervals (c_{2j}, c_{2j+1}). If two consecutive elements of $\{c_0, \dots, c_{2n+1}\}$ collide, then $\omega(\lambda)$ is, a priori, not well-defined. Thus, it is natural to ask: How does $\omega(\lambda)$ behave at these collisions?

In the previous theorem we have solved this question at the set $\Lambda^{n-1} = G \cup R \cup S$, which covers just the geodesic flow limit: $c_1 \to 0^+$, the n simple regular collapses: $c_{2l+1}, c_{2l} \to c^*$ for some l, and the n simple singular collapses: $c_{2l-1}, c_{2l} \to c^*$ for some l. But there are many

more (multiple) collapses, from double ones to total ones. Double collapses correspond to the set Λ^{n-2}. Total collapses have multiplicity n, so they correspond to set of vertexes Λ^0.

We believe that it does not make sense to describe the asymptotic behaviour of the frequency map at all of them, since the behaviour in each case must be the expected one. In order to convince the reader of the validity of this claim, we end the paper with a couple of extreme cases.

As a first example, let us consider the vertex $\hat{\lambda} = (a_1, \dots, a_n) \in \Lambda^0$. It represents the unique total singular collapse, because it is the unique common vertex of the 2^n open connected components of the caustic space:

$$\bigcap_{\sigma\in\{0,1\}^n} \bar{\Lambda}_\sigma = \{\hat{\lambda}\} = \bigcap_{j=1}^{n} \bar{S}_j.$$

Using that the point $\hat{\lambda}$ belongs to all the closures $\bar{S}_j$, from theorem 14 we get that $\omega(\hat{\lambda}) = (1/2, \dots, 1/2)$. Which is the asymptotic behaviour of ω at this vertex? In Appendix A.7 it is proved that

$$\omega(\lambda) = \omega(\hat{\lambda}) + \mathrm{O}(1/\log|a_1 - \lambda_1|, \dots, 1/\log|a_n - \lambda_n|), \qquad \lambda \to \hat{\lambda}.$$

This behaviour is singular in the n caustic coordinates, as expected.

On the contrary, the vertex $\tilde{\lambda} = (a_2, \dots, a_{n+1}) \in \Lambda^0$ represents the unique total regular collapse, so we predict a regular behaviour in the n caustic coordinates. In Appendix A.6 we show that $\omega(\tilde{\lambda}) = (\widetilde{\omega}_1, \dots, \widetilde{\omega}_n)$, where the limit frequencies $0 < \widetilde{\omega}_j < 1/2$ are defined as $\sin^2 \pi\widetilde{\omega}_j = a_1/a_{j+1}$, and the asymptotic behaviour is

$$\omega(\lambda) = \omega(\tilde{\lambda}) + \mathrm{O}(\tilde{\lambda} - \lambda), \qquad \lambda \to \tilde{\lambda}^-.$$

Once more, the frequency map has the expected behaviour.

7. Conclusion and further questions

We studied periodic trajectories of billiards inside nondegenerate ellipsoids of $\mathbb{R}^{n+1}$. First, we trivially extended the definition of the frequency map ω to any dimension, we presented four conjectures about ω based on numerical computations, and we deduced from the last one some lower bounds on the periods. Next, we proved that ω can be continuously extended to any singular value of the caustic parameters, although it is exponentially sharp at the "inner" singular caustic parameters. Finally, we focused on ellipses and triaxial ellipsoids, where we found examples of trajectories whose periods coincide with the previous lower bounds. We also computed several bifurcation curves. Despite these results, many unsolved questions remain. We indicate just four.

The most obvious challenge is to tackle any of the conjectures, although it does not look easy. We have already devoted some efforts without success. We believe that the proof of any of these conjectures requires either a deep use of algebraic geometry or to rewrite the frequency map as the gradient of a "Hamiltonian"; see [43, §4].

Another interesting question is to describe completely the phase space of billiards inside ellipsoids in $\mathbb{R}^{n+1}$ for $n \geq 2$. A rich hierarchy of invariant objects appears in these billiards: Liouville maximal tori, low dimensional tori, normally hyperbolic manifolds whose stable and unstable manifolds are doubled, et cetera. For instance, the stable and unstable invariant manifolds of the two-periodic hyperbolic trajectory corresponding to an oscillation along the major axis of the ellipsoid were fully described in [13].

Third, we plan to give a complete classification of the symmetric periodic trajectories inside generic ellipsoids [9]. To present the problem, let us consider the symmetric periodic trajectories inside an ellipse displayed in figure 3. On the one hand, the three-periodic trajectory drawn in a continuous red line has an impact point on (and is symmetric with respect to) the x-axis. On the other hand, the four-periodic trajectory drawn in a dashed green line has a couple of segments passing through (and is symmetric with respect to) the origin. It is immediate to realize that there do not exist neither a trajectory with a hyperbola as caustic like the first one, neither a trajectory with an ellipse as caustic like the second one. The problem consists of describing all possible kinds of symmetric periodic trajectories once fixed the type of the n caustics for ellipsoids in $\mathbb{R}^{n+1}$. Once these trajectories were well understood, we could study their persistence under small symmetric perturbations of the ellipsoid, and the break-up of the Liouville tori on which they live. Similar results have already been found in other billiard frameworks: homoclinic trajectories inside ellipsoids of $\mathbb{R}^{n+1}$ with a unique major axis [8], and periodic trajectories inside circumferences of the plane [34].

Finally, we look for simple formulae to express the caustic parameters $\lambda_1, \ldots, \lambda_n$ that give rise to periodic trajectories of small periods in terms of the parameters $a_1, \ldots, a_{n+1}$ of the ellipsoid. As a by-product of the those formulae, one can find algebraic expressions for some of the bifurcation curves displayed in subsection 5.4. This is a work in progress [35].

Acknowledgments

P. S. Casas was supported in part by MCyT-FEDER Grant MTM2006-00478 (Spain). R. Ramírez-Ros was supported in part by MICINN-FEDER Grant MTM2009-06973 (Spain) and CUR-DIUE Grant 2009SGR859 (Catalonia). Useful conversations with Jaume Amorós, Àlex Haro, Yuri Fedorov and Carles Simó are gratefully acknowledged.

Appendix A. Computations with hyperelliptic integrals

Appendix A.1. Technical lemmas

Lemma 15. *Let $f_\epsilon \in C^0([\alpha, \beta])$ be a family of functions such that $f_\epsilon = f_0 + \mathrm{O}(\epsilon)$ in the C^0-topology. Then*

$$I_\epsilon = \int_\alpha^\beta \frac{f_\epsilon(s)\mathrm{d}s}{\sqrt{(s-\alpha)(\beta-s)}} = \int_\alpha^\beta \frac{f_0(s)\mathrm{d}s}{\sqrt{(s-\alpha)(\beta-s)}} + \mathrm{O}(\epsilon).$$

Proof. $|I_\epsilon - I_0| \leq |f_\epsilon - f_0|_{C^0([\alpha,\beta])} \int_\alpha^\beta ((s-\alpha)(\beta-s))^{-1/2}\mathrm{d}s = \pi |f_\epsilon - f_0|_{C^0([\alpha,\beta])} = \mathrm{O}(\epsilon)$. □

Lemma 16. *Let $f \in C^1([m,M])$ with $m < \alpha < \beta < M$ and $\epsilon = \beta - \alpha$. Then*

$$\int_\alpha^\beta \frac{f(s)\mathrm{d}s}{\sqrt{(s-\alpha)(\beta-s)}} = \pi f(\alpha) + \mathrm{O}(\epsilon) = \pi f(\beta) + \mathrm{O}(\epsilon), \qquad \epsilon \to 0^+.$$

Proof. Using the Mean Value Theorem for integrals, we get that there exists some $s_0 \in [\alpha,\beta]$ such that the integral is equal to $f(s_0)\int_\alpha^\beta ((s-\alpha)(\beta-s))^{-1/2}\mathrm{d}s = \pi f(s_0)$. □

Lemma 17. *Let $f \in C^1([0,M])$ with $0 < \epsilon < M$. Then*

$$I_\epsilon = \int_0^\epsilon \frac{f(s)\mathrm{d}s}{\sqrt{\epsilon - s}} = 2f(0)\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2}), \qquad \epsilon \to 0^+.$$

Proof. $I_\epsilon = \left[-2(\epsilon-s)^{1/2}f(s)\right]_{s=0}^{s=\epsilon} + 2\int_0^\epsilon (\epsilon-s)^{1/2}f'(s)\mathrm{d}s = 2f(0)\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2})$. □

Lemma 18. *Let $f \in C^1([\alpha,\beta])$. Set $\eta = f(\alpha)\log(4\beta-4\alpha) + \int_\alpha^\beta (s-\alpha)^{-1}(f(s)-f(\alpha))\mathrm{d}s$, $\xi = \int_\alpha^\beta (s-\alpha)^{-3/2}(f(s)-f(\alpha))\mathrm{d}s$, $\mu = f(\beta)\log(4\beta-4\alpha) + \int_\alpha^\beta (\beta-s)^{-1}(f(s)-f(\beta))\mathrm{d}s$, and $\psi = \int_\alpha^\beta (\beta-s)^{-3/2}(f(s)-f(\beta))\mathrm{d}s$. Then*

$$\int_\alpha^\beta \frac{f(s)\mathrm{d}s}{\sqrt{(s+\epsilon-\alpha)(s-\alpha)}} = -f(\alpha)\log\epsilon + \eta + \mathrm{O}(\epsilon\log\epsilon), \qquad \epsilon \to 0^+,$$

$$\int_\alpha^\beta \frac{f(s)\mathrm{d}s}{(s+\epsilon-\alpha)\sqrt{s-\alpha}} = \pi f(\alpha)\epsilon^{-1/2} + \xi + \mathrm{O}(\epsilon^{1/2}), \qquad \epsilon \to 0^+,$$

$$\int_\alpha^\beta \frac{f(s)\mathrm{d}s}{\sqrt{(\beta+\epsilon-s)(\beta-s)}} = -f(\beta)\log\epsilon + \mu + \mathrm{O}(\epsilon\log\epsilon), \qquad \epsilon \to 0^+,$$

$$\int_\alpha^\beta \frac{f(s)\mathrm{d}s}{(\beta+\epsilon-s)\sqrt{\beta-s}} = \pi f(\beta)\epsilon^{-1/2} + \psi + \mathrm{O}(\epsilon^{1/2}), \qquad \epsilon \to 0^+.$$

The first (respectively, last) two estimates also hold when f has a singularity at $s = \beta$ (respectively, at $s = \alpha$), provided $f \in L^1([\alpha,\beta])$.

Proof. We split the first integral as $I_\epsilon = \tilde\eta + \hat I_\epsilon - \tilde I_\epsilon$, where $\tilde\eta = \int_\alpha^\beta (s-\alpha)^{-1}(f(s)-f(\alpha))\mathrm{d}s$ is a constant, and

$$\hat I_\epsilon = \int_\alpha^\beta \frac{f(\alpha)\mathrm{d}s}{\sqrt{(s+\epsilon-\alpha)(s-\alpha)}}, \qquad \tilde I_\epsilon = \int_\alpha^\beta \frac{f(s)-f(\alpha)}{s-\alpha}\left(1 - \sqrt{\frac{s-\alpha}{s+\epsilon-\alpha}}\right)\mathrm{d}s.$$

By performing the change $x^2 = s - \alpha$ in the integral $\hat I_\epsilon$, we get that

$$\hat I_\epsilon = 2\int_0^{\sqrt{\beta-\alpha}} \frac{f(\alpha)\mathrm{d}x}{\sqrt{x^2+\epsilon}} = 2f(\alpha)\left[\log\left(x + \sqrt{x^2+\epsilon}\right)\right]_{x=0}^{x=\sqrt{\beta-\alpha}} = -f(\alpha)\log\epsilon + \hat\eta + \mathrm{O}(\epsilon),$$

where $\hat\eta = f(\alpha)\log(4\beta-4\alpha)$ is another constant. Thus, to get the first formula with constant $\eta = \hat\eta + \tilde\eta$ it suffices to see that $\tilde I_\epsilon = \mathrm{O}(\epsilon\log\epsilon)$.

Once fixed some $\gamma \in (\alpha,\beta)$, we decompose the integral $\tilde I_\epsilon$ as the sum $\tilde J_\epsilon + \tilde K_\epsilon$, where $\tilde J_\epsilon = \int_\alpha^\gamma \tilde f(s) r_\epsilon(s)\mathrm{d}s$, $\tilde K_\epsilon = \int_\gamma^\beta \tilde f(s) r_\epsilon(s)\mathrm{d}s$, and

$$\tilde f(s) = \frac{f(s)-f(\alpha)}{s-\alpha}, \qquad r_\epsilon(s) = 1 - \sqrt{\frac{s-\alpha}{s+\epsilon-\alpha}}.$$

First, we consider the interval $[\alpha,\gamma]$. Then $|\tilde f|_\infty=\max\{|\tilde f(s)|:\alpha\le s\le\gamma\}<\infty$ and $r_\epsilon(s)$ is positive in $[\alpha,\gamma]$. Set $\delta=\gamma-\alpha$. Using again the change $x^2=s-\alpha$, we see that

$$\begin{aligned}|\tilde f|_\infty^{-1}|\tilde J_\epsilon|\le&\int_\alpha^\gamma r_\epsilon(s)\mathrm{d}s=\delta-\int_\alpha^\gamma\sqrt{\frac{s-\alpha}{s+\epsilon-\alpha}}\mathrm{d}s=\delta-2\int_0^{\sqrt\delta}\frac{x^2\mathrm{d}x}{\sqrt{x^2+\epsilon}}\\&=\delta-\left[x\sqrt{x^2+\epsilon}+\epsilon\log\left(x+\sqrt{x^2+\epsilon}\right)\right]_{x=0}^{x=\sqrt\delta}=-\frac{\epsilon}{2}\log\epsilon+\mathrm{O}(\epsilon)=\mathrm{O}(\epsilon\log\epsilon).\end{aligned}$$

Concerning the other interval, we note that $r_\epsilon(s)$ is positive and decreasing in $[\gamma,\beta]$. Hence, $\max\{|r_\epsilon(s)|:\gamma\le s\le\beta\}=r_\epsilon(\gamma)$, and so $|\tilde K_\epsilon|\le r_\epsilon(\gamma)\int_\gamma^\beta|\tilde f(s)|\mathrm{d}s=\mathrm{O}(\epsilon)$. This ends the proof of the first formula.

We split the second integral as $L_\epsilon=\xi+\hat L_\epsilon-\tilde L_\epsilon$, where ξ is the constant given in the statement of the lemma, and

$$\hat L_\epsilon=\int_\alpha^\beta\frac{f(\alpha)\mathrm{d}s}{(s+\epsilon-\alpha)\sqrt{s-\alpha}},\qquad\tilde L_\epsilon=\int_\alpha^\beta\frac{f(s)-f(\alpha)}{(s-\alpha)^{3/2}}\left(1-\frac{s-\alpha}{s+\epsilon-\alpha}\right)\mathrm{d}s.$$

By performing the change $x=s-\alpha$ in the integral $\hat L_\epsilon$, we get that

$$\hat L_\epsilon=\int_0^{\beta-\alpha}\frac{f(\alpha)\mathrm{d}x}{(x+\epsilon)\sqrt{x}}=2f(\alpha)\epsilon^{-1/2}\left[\operatorname{atan}\left(\epsilon^{-1/2}x^{1/2}\right)\right]_{x=0}^{x=\beta-\alpha}=\pi f(\alpha)\epsilon^{-1/2}+\mathrm{O}(\epsilon^{1/2}).$$

Thus, to get the second formula it suffices to see that $\tilde L_\epsilon=\mathrm{O}(\epsilon^{1/2})$, which follows from similar computations than the ones above.

The last formulae are obtained by performing the change of variables $s-\alpha=\beta-t$ in the former ones. □

Corollary 19. *Let $f\in C^1([m,M])$ with $m<\alpha_-<\alpha_+<\beta_-<\beta_+<M$, and*

$$I=I(\alpha_-,\alpha_+,\beta_-,\beta_+)=\int_{\alpha_+}^{\beta_-}\frac{f(s)\mathrm{d}s}{\sqrt{(s-\alpha_-)(s-\alpha_+)(\beta_--s)(\beta_+-s)}}.$$

Let α_* and β_* be two reals such that $m<\alpha_*<\beta_*<M$. Let $\epsilon=(\epsilon_1,\epsilon_2)\in\mathbb{R}_+^2$ with $\epsilon_1=\alpha_+-\alpha_-$ and $\epsilon_2=\beta_+-\beta_-$. Then there exists a constant $\zeta\in\mathbb{R}$ such that

$$I=-\frac{f(\alpha_*)(1+\mathrm{O}(\epsilon_2))\log\epsilon_1+f(\beta_*)(1+\mathrm{O}(\epsilon_1))\log\epsilon_2}{\beta_*-\alpha_*}+\zeta+\mathrm{O}(\epsilon_1\log\epsilon_1,\epsilon_2\log\epsilon_2),$$

as $\alpha_\pm\to\alpha_*$ and $\beta_\pm\to\beta_*$, so that $\epsilon=(\epsilon_1,\epsilon_2)\to(0^+,0^+)$.

Proof. It follows by applying the first and third estimates of the previous lemma to the integrals $\int_{\alpha_+}^\gamma$ and $\int_\gamma^{\beta_-}$ for some point $\gamma\in(\alpha_+,\beta_-)$, although before we must fix the lower limit of the first integral with the change $x-\alpha_*=s-\alpha_+$, and the upper limit of the second integral with the change $x-\beta_*=s-\beta_-$. □

Lemma 20. *Let $\mathbf{K}_\epsilon\omega_\epsilon=\tau_\epsilon$ be a family of square linear systems defined for $\epsilon>0$.*

(i) If the limits $\mathbf{K}=\lim_{\epsilon\to0^+}\mathbf{K}_\epsilon$ and $\tau=\lim_{\epsilon\to0^+}\tau_\epsilon$ exist, and $\mathbf{K}$ is nonsingular, then

$$\omega_\epsilon=\omega+\mathrm{O}(|\mathbf{K}_\epsilon-\mathbf{K}|,|\tau_\epsilon-\tau|),\qquad\epsilon\to0^+,$$

where $\omega=\mathbf{K}^{-1}\tau$ is the unique solution of the nonsingular limit system $\mathbf{K}\omega=\tau$.

(ii) If, in addition, the matrix $\mathbf{K}_\epsilon$ *and the vector* τ_ϵ *are differentiable at* $\epsilon = 0$*, then the solution* ω_ϵ *also is differentiable at* $\epsilon = 0$*. To be more precise, if*

$$\mathbf{K}_\epsilon = \mathbf{K} + \epsilon\mathbf{L} + \mathrm{o}(\epsilon), \qquad \tau_\epsilon = \tau + \epsilon\zeta + \mathrm{o}(\epsilon), \qquad \epsilon \to 0^+,$$

for some square matrix $\mathbf{L}$ *and some vector* ζ*, then*

$$\omega_\epsilon = \omega + \epsilon\kappa + \mathrm{o}(\epsilon), \qquad \epsilon \to 0^+,$$

where $\omega = \mathbf{K}^{-1}\tau$ *and* $\kappa = \mathbf{K}^{-1}(\zeta - \mathbf{L}\omega)$.

Proof. Both results follow directly from classical error bounds in numerical linear algebra. See, for instance, [23, §2.7]. □

Appendix A.2. Another characterization of the frequency

We associate a "frequency" $\omega = \varpi(c) \in \mathbb{R}^n$ to any vector $c = (c_1, \ldots, c_{2n+1}) \in \mathbb{R}^{2n+1}$ such that $c_0 := 0 < c_1 < \cdots < c_{2n+1}$ in the following way. First, we consider:

- The polynomial $T(s) = \prod_{i=1}^{2n+1}(c_i - s) \in \mathbb{R}_{2n+1}[s]$, which is positive in the $n+1$ intervals of the form (c_{2j}, c_{2j+1});
- The $n+1$ linear functionals $P(s) \mapsto \mathcal{K}_j[P(s)] = \int_{c_{2j}}^{c_{2j+1}} (T(s))^{-1/2} P(s)\mathrm{d}s$;
- The $n+1$ column vectors $K_j = (\mathcal{K}_j[1], \mathcal{K}_j[s], \ldots, \mathcal{K}_j[s^{n-1}])^t \in \mathbb{R}^n$;
- The $n \times n$ nonsingular matrix $\mathbf{K} = (-K_1, \ldots, (-1)^n K_n)$; and
- The linear functionals $\mathcal{K}[P(s); \omega] = \mathcal{K}_0[P(s)] + 2\sum_{j=1}^n (-1)^j \omega_j \mathcal{K}_j[P(s)]$, for $\omega \in \mathbb{R}^n$.

The hypothesis $c_1 > 0$ is not essential to get a nonsingular matrix $\mathbf{K}$, but it suffices to assume the strict inequalities $c_1 < \cdots < c_{2n+1}$; see [24, §III.3].

Lemma 21. *There exists an unique* $\omega \in \mathbb{R}^n$ *such that*

$$\mathcal{K}[P(s); \omega] = 0, \qquad \forall P(s) \in \mathbb{R}_{n-1}[s], \tag{A.1}$$

or equivalently, such that $K_0 + 2\mathbf{K}\omega = 0$*, which is the matricial form of the linear system given in (8).*

Proof. By taking the basis $\{1, s, \ldots, s^{n-1}\}$ of $\mathbb{R}_{n-1}[s]$, we see that condition (A.1) is equivalent to the linear system $K_0 + 2\mathbf{K}\omega = 0$. □

Therefore, condition (A.1) is an equivalent characterization of the frequency. From now on, $\omega = \varpi(c)$ stands for the frequency computed through the previous steps.

Appendix A.3. Geodesic flow limit: $c_1 \to 0^+$

Let $K_{00}^G = 2(\prod_{i=2}^n c_i)^{-1/2}$, $K_0^G = (K_{00}^G, 0, \ldots, 0) \in \mathbb{R}^n$, and $T^G(s) = -s\prod_{i=2}^{2n+1}(c_i - s)$. Let $\mathbf{K}^G$ be the $n \times n$ nonsingular matrix associated to the vector $c^G = (0, c_2, \ldots, c_{2n+1})$. Let $\kappa^G \in \mathbb{R}^n$ be the unique solution of the linear system $K_0^G + 2\mathbf{K}^G\kappa^G = 0$. Then

$$\omega = \kappa^G c_1^{1/2} + \mathrm{O}(c_1^{3/2}), \qquad c_1 \to 0^+. \tag{A.2}$$

The proof is short. First, we note that $T = T^G + \mathrm{O}(c_1)$ uniformly in $[0, c_{2n+1}]$. Thus, using lemma 15, we get that $\mathbf{K} = \mathbf{K}^G + \mathrm{O}(c_1)$ as $c_1 \to 0^+$. And using lemma 17 we see that $K_0 = K_0^G c_1^{1/2} + \mathrm{O}(c_1^{3/2})$ as $c_1 \to 0^+$. Therefore, the linear systems $K_0^G + 2\mathbf{K}^G \kappa^G = 0$ and $c_1^{-1/2} K_0 + 2\mathbf{K}(c_1^{-1/2}\omega) = 0$ are $\mathrm{O}(c_1)$-close, being $\mathbf{K}^G$ nonsingular, so (A.2) follows from the first item in lemma 20.

Appendix A.4. Simple regular collapse: $c_{2l+1}, c_{2l} \to c^*$ for some $l = 1, \ldots, n$

Set $c^R = (c_1, \ldots, c_{2l-1}, c_{2l+2}, \ldots, c_{2n+1}) \in \mathbb{R}^{2n-1}$. Let $T^R(s) = \prod_{i \neq 2l, 2l+1}(c_i - s)$ be the polynomial associated to c^R. Let $\mathcal{K}_j^R$ and $\mathcal{K}^R$ be the functionals associated to c^R. Let $\omega^R = (\omega_1^R, \ldots, \omega_n^R) \in \mathbb{R}^n$, where $\omega_{\neq l}^R := (\omega_1^R, \ldots, \omega_{l-1}^R, \omega_{l+1}^R, \ldots, \omega_n^R) = \varpi(c^R) \in \mathbb{R}^{n-1}$ is the frequency associated to c^R, and $\omega_l^R \in \mathbb{R}$ is determined by

$$\int_0^{c_1} \frac{\mathrm{d}s}{|c^* - s|\sqrt{T^R(s)}} + 2\sum_{j \neq l} \int_{c_{2j}}^{c_{2j+1}} \frac{(-1)^j \omega_j^R \mathrm{d}s}{|c^* - s|\sqrt{T^R(s)}} + \frac{(-1)^l 2\pi \omega_l^R}{\sqrt{-T^R(c^*)}} = 0. \tag{A.3}$$

Let $\epsilon = c_{2l+1} - c_{2l}$. Then

$$\omega = \omega^R + \mathrm{O}(\epsilon), \qquad c_{2l+1}, c_{2l} \to c^*. \tag{A.4}$$

In order to prove this claim, we first observe that characterization (A.1) is equivalent to the system of n linear equations

$$\begin{cases} \mathcal{K}[(c^* - s)s^i; \omega] = 0 \text{ for } i = 0, \ldots, n-2 \\ \mathcal{K}[1; \omega] = 0 \end{cases}, \tag{A.5}$$

because $\{1, c^* - s, \ldots, (c^* - s)s^{n-2}\}$ is a basis of $\mathbb{R}_{n-1}[s]$. Now, using lemmas 15 and 16, we deduce the estimates

$$\mathcal{K}_j[(c^* - s)s^i] = \int_{c_{2j}}^{c_{2j+1}} \frac{(c^* - s)s^i \mathrm{d}s}{|c^* - s|\sqrt{T^S(s)} + \mathrm{O}(\epsilon)} = \begin{cases} \mathcal{K}_j^R[s^i] + \mathrm{O}(\epsilon) & \text{if } j < l, \\ \mathrm{O}(\epsilon) & \text{if } j = l, \\ -\mathcal{K}_{j-1}^R[s^i] + \mathrm{O}(\epsilon) & \text{if } j > l; \end{cases}$$

and

$$\mathcal{K}_j[1] = \begin{cases} \pi \left(-T^R(c^*)\right)^{-1/2} + \mathrm{O}(\epsilon) & \text{if } j = l, \\ \displaystyle\int_{c_{2j}}^{c_{2j+1}} \frac{\mathrm{d}s}{|c^* - s|\sqrt{T^R(s)}} + \mathrm{O}(\epsilon) & \text{otherwise.} \end{cases}$$

Therefore, the linear system (A.5) is $\mathrm{O}(\epsilon)$-close to the nonsingular linear system

$$\begin{cases} \mathcal{K}^R[s^i; \omega_{\neq l}^R] = 0 \text{ for } i = 0, \ldots, n-2 \\ \text{condition (A.3)} \end{cases},$$

and the asymptotic formula (A.4) follows from the first item in lemma 20.

Appendix A.5. Simple singular collapse: $c_{2l-1}, c_{2l} \to c^*$ for some $l = 1, \dots, n$

Set $c^S = (c_1, \dots, c_{2l-2}, c_{2l+1}, \dots, c_{2n+1}) \in \mathbb{R}^{2n-1}$. Let $T^S(s) = \prod_{i \neq 2l-1, 2l}(c_i - s)$ be the polynomial associated to c^S. Let $\mathcal{K}_j^S$ and $\mathcal{K}^S$ be the functionals associated to c^S. Let $\omega^S = (\omega_1^S, \dots, \omega_n^S) \in \mathbb{R}^n$, where $\omega_{\neq l}^S := (\omega_1^S, \dots, \omega_{l-1}^S, \omega_{l+1}^S, \dots, \omega_n^S) = \varpi(c^S) \in \mathbb{R}^{n-1}$ and

$$\omega_l^S = \begin{cases} 1/2 & \text{if } l = 1 \\ \omega_{l-1}^S & \text{otherwise} \end{cases}.$$

Let $\epsilon = c_{2l} - c_{2l-1} > 0$ and $\delta = |\log \epsilon|^{-1} > 0$. Then there exists $\kappa^S \in \mathbb{R}^n$ such that

$$\omega = \omega^S + \delta \kappa^S + \mathrm{o}(\delta), \qquad c_{2l-1}, c_{2l} \to c^*. \tag{A.6}$$

To prove this claim, we set $d = \sqrt{T^S(c^*)} > 0$. We know that characterization (A.1) is equivalent to the system of n linear equations

$$\begin{cases} \mathcal{K}[\delta d; \omega] = 0 \\ \mathcal{K}[(c^* - s)s^i; \omega] = 0 \text{ for } i = 0, \dots, n-2 \end{cases}, \tag{A.7}$$

because $\{\delta d, c^* - s, \dots, (c^* - s)s^{n-2}\}$ is a basis of $\mathbb{R}_{n-1}[s]$. Now, using lemmas 15 and 18, we deduce the following asymptotic estimates. On the one hand, there exist some constants $\zeta_0, \zeta_1, \dots, \zeta_n \in \mathbb{R}$ such that

$$\mathcal{K}_j[\delta d] = \delta d \mathcal{K}_j[1] = \begin{cases} 1 + \zeta_j \delta + \mathrm{O}(\epsilon) & \text{if } j = l-1, l, \\ \zeta_j \delta + \mathrm{O}(\epsilon\delta) & \text{otherwise.} \end{cases}$$

On the other hand,

$$\mathcal{K}_j[(c^* - s)s^i] = \int_{c_{2j}}^{c_{2j+1}} \frac{(c^* - s)s^i \mathrm{d}s}{|c^* - s|\sqrt{T^S(s) + \mathrm{O}(\epsilon)}} = \begin{cases} \mathcal{K}_j^S[s^i] + \mathrm{O}(\epsilon) & \text{if } j < l-1, \\ \int_{c_{2l-2}}^{c^*} \frac{s^i \mathrm{d}s}{\sqrt{T^S(s)}} + \mathrm{O}(\epsilon) & \text{if } j = l-1, \\ -\int_{c^*}^{c_{2l+1}} \frac{s^i \mathrm{d}s}{\sqrt{T^S(s)}} + \mathrm{O}(\epsilon) & \text{if } j = l, \\ -\mathcal{K}_{j-1}^S[s^i] + \mathrm{O}(\epsilon) & \text{if } j > l. \end{cases}$$

In particular, $\mathcal{K}_{l-1}[(c^* - s)s^i] - \mathcal{K}_l[(c^* - s)s^i] = \int_{c_{2l-2}}^{c_{2l+1}} \frac{s^i \mathrm{d}s}{\sqrt{T^S(s)}} + \mathrm{O}(\epsilon) = \mathcal{K}_{l-1}^S[s^i] + \mathrm{O}(\epsilon)$.

We assume now that $l \neq 1$. The case $l = 1$ is studied later on. Since $\epsilon \ll \delta$, the linear system (A.7) is $\mathrm{O}(\delta)$-close to the nonsingular linear system

$$\begin{cases} 2(-1)^{l-1}(\omega_{l-1}^S - \omega_l^S) = 0 \\ \mathcal{K}^S[s^i; \omega_{\neq l}^S] = 2(-1)^l(\omega_l^S - \omega_{l-1}^S) \int_{c^*}^{c_{2l+1}} \frac{s^i \mathrm{d}s}{\sqrt{T^S(s)}} \text{ for } i = 0, \dots, n-2 \end{cases},$$

which in its turn is equivalent to the linear system

$$\begin{cases} \omega_l^S = \omega_{l-1}^S \\ \mathcal{K}^S[s^i; \omega_{\neq l}^S] = 0 \text{ for } i = 0, \dots, n-2 \end{cases} \tag{A.8}$$

whose unique solution is $\omega_{\neq l}^S = \varpi(c^S)$ and $\omega_l^S = \omega_{l-1}^S$.

Thus, the asymptotic formula $\omega = \omega^S + \mathrm{O}(\delta)$ follows from the first item in lemma 20. In fact, this result can be improved using the second item in lemma 20. It suffices to note that the linear system (A.7) is not only $\mathrm{O}(\delta)$-equivalent to (A.8), but it is differentiable at $\delta = 0$.

Hence, (A.6) holds for some vector κ^S that could be explicitely computed in terms of the limit system and the constants $\zeta_0, \dots, \zeta_n$.

If $l = 1$, the linear system (A.7) is $\mathrm{O}(\delta)$-equivalent to the nonsingular linear system

$$\begin{cases} \omega_1^S = 1/2 \\ \mathcal{K}^S[s^i; \omega_{\neq 1}^S] = 0 \text{ for } i = 0, \dots, n-2 \end{cases},$$

and the proof ends with just the same arguments that for $l \neq 1$. We omit the details.

Appendix A.6. Total regular collapse: $c_{2l+1}, c_{2l} \to c_l^*$ for all $l = 1, \dots, n$

Let us study the case of n simultaneous collapses, all of them regular. That is, once fixed a vector $c^* = (c_1^*, \dots, c_n^*) \in \mathbb{R}^n$ such that $0 < c_1 < c_1^* < \cdots < c_n^*$, we study the asymptotic behaviour of the frequency $\omega = \varpi(c)$ when $c_{2l+1}, c_{2l} \to c_l^*$ for all $l = 1, \dots, n$. Let $\widetilde{\omega} = (\widetilde{\omega}_1, \dots, \widetilde{\omega}_n) \in \mathbb{R}^n$ be the vector whose components verify that $0 < \widetilde{\omega}_l < 1/2$ and $\sin^2 \pi\widetilde{\omega}_l = c_1/c_l^*$. Let $\epsilon = (\epsilon_1, \dots, \epsilon_n) \in \mathbb{R}_+^n$ with $\epsilon_l = c_{2l+1} - c_{2l}$. Then

$$\omega = \widetilde{\omega} + \mathrm{O}(\epsilon), \qquad \epsilon \to (0^+, \dots, 0^+). \tag{A.9}$$

Let $Q_l = \sqrt{c_l^* - c_1} \prod_{i \neq l} |c_i^* - c_l^*| > 0$. Let $\{P_1(s), \dots, P_n(s)\}$ be the basis of $\mathbb{R}_{n-1}[s]$ univocally determined by the interpolating conditions

$$P_l(c_j^*) = \begin{cases} Q_l & \text{if } j = l, \\ 0 & \text{otherwise.} \end{cases}$$

That is, $P_l(s) = (-1)^{l-1} \sqrt{c_l^* - c_1} \prod_{i \neq l} (c_i^* - s)$. Using lemma 16, we get the estimates

$$\mathcal{K}_0[P_l(s)] = \int_0^{c_1} \left(\frac{(-1)^{l-1} \sqrt{c_l^* - c_1}}{(c_l^* - s)\sqrt{c_1 - s}} + \mathrm{O}(\epsilon) \right) \mathrm{d}s = 2(-1)^{l-1} \operatorname{atan} \sqrt{\frac{c_1}{c_l^* - c_1}} + \mathrm{O}(\epsilon),$$

$\mathcal{K}_l[P_l(s)] = \pi + \mathrm{O}(\epsilon)$, and $\mathcal{K}_j[P_l(s)] = \mathrm{O}(\epsilon)$ for $j \neq 0, l$. Thus, the $n \times n$ linear system

$$\mathcal{K}[P_l(s); \omega] = 0 \text{ for } l = 1, \dots, n$$

is $\mathrm{O}(\epsilon)$-close to the nonsingular decoupled linear system

$$2(-1)^{l-1} \left(\operatorname{atan} \sqrt{\frac{c_1}{c_l^* - c_1}} - \pi\widetilde{\omega}_l \right) = 0 \text{ for } l = 1, \dots, n,$$

whose unique solution is given by $\tan^2 \pi\widetilde{\omega}_l = c_1/(c_l^* - c_1)$, and so, by $\sin^2 \pi\widetilde{\omega}_l = c_1/c_l^*$. Hence, the asymptotic formula (A.9) follows from the first item in lemma 20.

Appendix A.7. Total singular collapse: $c_{2l-1}, c_{2l} \to c_l^*$ for all $l = 1, \dots, n$

Let $\widehat{\omega} = (1/2, \dots, 1/2) \in \mathbb{R}^n$, $\epsilon = (\epsilon_1, \dots, \epsilon_n) \in \mathbb{R}_+^n$, and $\delta = (\delta_1, \dots, \delta_n) \in \mathbb{R}_+^n$, where $\epsilon_l = c_{2l} - c_{2l-1}$ and $\delta_l = |\log \epsilon_l|^{-1}$. Then

$$\omega = \widehat{\omega} + \mathrm{O}(\delta), \qquad \epsilon \to (0^+, \dots, 0^+). \tag{A.10}$$

Remark 12. By applying repeatedly the result on simple singular collapses, we see that

$$\lim_{\epsilon_n\to 0^+}\left(\cdots\lim_{\epsilon_2\to 0^+}\left(\lim_{\epsilon_1\to 0^+}\omega\right)\right)=\widehat{\omega}.$$

In fact, these repeated limits can be taken in any order. Nevertheless, this result is weaker than estimate (A.10), so we need a formal proof of the estimate.

We consider the constants $Q_l=\sqrt{c_{2n+1}-c_l^*}\prod_{i\neq l}|c_i^*-c_l^*|$. Let $\{P_1(s),\dots,P_n(s)\}$ be the basis of $\mathbb{R}_{n-1}[s]$ univocally determined by

$$P_l(c_j^*)=\begin{cases}Q_l & \text{if } j=l,\\ 0 & \text{otherwise.}\end{cases}$$

Now, using corollary 19, we get that there exists some constants $\zeta_{jl}\in\mathbb{R}$ such that

$$\mathcal{K}_j[\delta_l P_l(s)]=\delta_l\mathcal{K}_j[P_l(s)]=\begin{cases}1+\zeta_{ll}\delta_l+\mathrm{o}(\delta) & \text{if } j=l,\\ \zeta_{jl}\delta_l+\mathrm{o}(\delta) & \text{otherwise,}\end{cases}$$

where $0\leq j\leq n$ and $1\leq l\leq n$. Therefore, the $n\times n$ linear system

$$\delta_l\mathcal{K}[P_l(s);\omega]=0 \text{ for } l=1,\dots,n$$

is $\mathrm{O}(\delta)$-close to the nonsingular linear system

$$\begin{cases}1-2\widehat{\omega}_1=0\\ 2(-1)^{l-1}(\widehat{\omega}_{l-1}-\widehat{\omega}_l)=0 \text{ for } l=2,\dots,n\end{cases}$$

whose unique solution is $\widehat{\omega}=(1/2,\dots,1/2)$. Thus, the asymptotic formula (A.10) follows from the first item in lemma 20.

Remark 13. The vectorial estimate (A.10) can be refined in several ways. For instance, one can get the componentwise estimates $\omega_1=1/2+\mathrm{O}(\delta_1)$ and $\omega_l=\omega_{l-1}+\mathrm{O}(\delta_l)$ for $l>1$. In particular, $\omega_l=1/2+\mathrm{O}(\delta_1,\dots,\delta_l)$. Even more, there exists a $n\times n$ constant *lower triangular* matrix $\mathbf{L}$ such that

$$\omega=\widehat{\omega}+\mathbf{L}\delta+\mathrm{o}(\delta),\qquad \epsilon\to(0^+,\dots,0^+).$$

We omit the proof, since we do not need this result and the computations are cumbersome.

Appendix A.8. Asymptotic behaviour of the function ν_{x}

The function $\nu_{\mathrm{x}}:(0,c)\cup(c,b)\to\mathbb{R}$ verifies that $I(\lambda)+J(\lambda)\rho_{\mathrm{x}}(\lambda)+K(\lambda)\nu_{\mathrm{x}}(\lambda)=0$, where the coefficients $I,J,K:(0,c)\cup(c,b)\to\mathbb{R}$ were given by

$$I(\lambda)=\int_0^{\underline{m}}\frac{\mathrm{d}s}{(a-s)\sqrt{T_{\mathrm{x}}(s)}},\quad J(\lambda)=-2\int_{\overline{m}}^{b}\frac{\mathrm{d}s}{(a-s)\sqrt{T_{\mathrm{x}}(s)}},\quad K(\lambda)=\frac{2\pi}{\sqrt{-T_{\mathrm{x}}(a)}},$$

with $T_{\mathrm{x}}(s)=(\lambda-s)(c-s)(b-s)$, $\underline{m}=\min(\lambda,c)$, and $\overline{m}=\max(\lambda,c)$. Here, $\rho_{\mathrm{x}}(\lambda)=\rho(\lambda;c,b)$ is the rotation function of the ellipse obtained by sectioning the ellipsoid Q with the coordinate plane $\{x=0\}$. The asymptotic properties of rotation functions of billiards inside ellipses were established in proposition 8.

First, let us consider the case $\epsilon:=\lambda\to 0^+$. Using lemmas 15 and 17 we get:

- $I(\epsilon) = I_0\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2})$, where $I_0 = 2a^{-1}(bc)^{-1/2}$;
- $J(\epsilon) = J_0 + \mathrm{O}(\epsilon)$, where $J_0 = -2\int_c^b (a-s)^{-1}(s(s-c)(b-s))^{-1/2}\mathrm{d}s$;
- $K(\epsilon) = K_0 + \mathrm{O}(\epsilon)$, where $K_0 = 2\pi(a(a-c)(a-b))^{-1/2}$;
- $\rho_{\mathrm{x}}(\epsilon) = \kappa^G\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2})$, where $\kappa^G = \kappa^G(b,c)$ can be found in proposition 8; and
- $\nu_{\mathrm{x}}(\epsilon) = -(I_0 + J_0\kappa^G)K_0^{-1}\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2}) = \mathrm{O}(\epsilon^{1/2})$. It is possible to check that $(I_0 + J_0\kappa^G)K_0^{-1} < 0$, but we do not need it.

Next, let us consider the case $\epsilon := b - \lambda \to 0^+$. We begin by computing the integral

$$r(\beta,\alpha) := \int_0^\beta \frac{\mathrm{d}s}{(\alpha-s)\sqrt{\beta-s}} = \frac{2}{\sqrt{\alpha-\beta}}\,\mathrm{atan}\sqrt{\beta/(\alpha-\beta)},$$

for any $0 < \beta < \alpha$. Then it is immediate to check that

$$\int_0^\beta \frac{\mathrm{d}s}{(\alpha_+-s)(\alpha_--s)\sqrt{\beta-s}} = \frac{r(\beta,\alpha_-) - r(\beta,\alpha_+)}{\alpha_+-\alpha_-},$$

for any $0 < \beta < \alpha_- < \alpha_+$. We also need the formula $r(\beta,\alpha) = 2\pi(\alpha-\beta)^{-1/2}\tilde{\rho}(\beta,\alpha)$, where $\tilde{\rho}(\beta,\alpha) := \lim_{\gamma\to\alpha^-}\rho(\gamma;\beta,\alpha)$ is one of the limits of the rotation number described in proposition 8. Using these formulae, jointly with lemmas 15 and 16, we see that:

- $I(b-\epsilon) = I_* + \mathrm{O}(\epsilon)$, where $I_* = 2\pi(a-b)^{-1}\left((b-c)^{-1/2}\tilde{\rho}(c,b) - (a-c)^{-1/2}\tilde{\rho}(c,a)\right)$;
- $J(b-\epsilon) = J_* + \mathrm{O}(\epsilon)$, where $J_* = -2\pi(a-b)^{-1}(b-c)^{-1/2}$;
- $K(b-\epsilon) = K_* + \mathrm{O}(\epsilon)$, where $K_* = 2\pi(a-b)^{-1}(a-c)^{-1/2}$;
- $\rho_{\mathrm{x}}(b-\epsilon) = \rho_{\mathrm{x}}(b) + \mathrm{O}(\epsilon) = \rho(b;c,b) + \mathrm{O}(\epsilon) = \tilde{\rho}(c,b) + \mathrm{O}(\epsilon)$; and
- $\nu_{\mathrm{x}}(b-\epsilon) = \tilde{\rho}(c,a) + \mathrm{O}(\epsilon) = \rho(a;c,a) + \mathrm{O}(\epsilon) = \rho_{\mathrm{y}}(a) + \mathrm{O}(\epsilon)$.

The estimates in the limit $\epsilon := \overline{m} - \underline{m} \to 0^+$, which equivals to $\lambda \to c$, are:

- $I(c\pm\epsilon) = -(a-c)^{-1}(b-c)^{-1/2}\log\epsilon + \mu + \mathrm{O}(\epsilon\log\epsilon)$, where μ is a constant that can be exactly computed from lemma 18;
- $J(c\pm\epsilon) = -(a-c)^{-1}(b-c)^{-1/2}\log\epsilon + \eta + \mathrm{O}(\epsilon\log\epsilon)$, where η is a constant that can be exactly computed from lemma 18;
- $K(c\pm\epsilon) = 2\pi(a-c)^{-1}(a-b)^{-1/2} + \mathrm{O}(\epsilon)$;
- $\rho_{\mathrm{x}}(c\pm\epsilon) = 1/2 + \kappa^S\log^{-1}\epsilon + \mathrm{O}(\log^{-2}\epsilon)$, where $\kappa^S = \kappa^S(c,b) = \mathrm{acosh}(b/c)^{1/2}$ according to proposition 8; and
- $\nu_{\mathrm{x}}(c\pm\epsilon) = (a-b)^{1/2}\left((a-c)(\eta-\mu) - 2(b-c)^{-1/2}\kappa^S\right)/2\pi + \mathrm{O}(\log^{-1}\epsilon)$. After some tedious, but simple, computations, one gets that $\nu_{\mathrm{x}}(c) = \tilde{\rho}(b,a) = \rho(a;b,a) = \rho_{\mathrm{z}}(a)$.

Appendix A.9. Asymptotic behaviour of the function ν_{y}

The function $\nu_{\mathrm{y}} : (b,a) \to \mathbb{R}$ verifies that $I(\lambda) + J(\lambda)\rho_{\mathrm{y}}(\lambda) + K(\lambda)\nu_{\mathrm{y}}(\lambda) = 0$, where the coefficients $I, J, K : (b,a) \to \mathbb{R}$ were given by

$$I(\lambda) = \int_0^c \frac{\mathrm{d}s}{(b-s)\sqrt{T_{\mathrm{y}}(s)}}, \quad J(\lambda) = 2\int_\lambda^a \frac{\mathrm{d}s}{(s-b)\sqrt{T_{\mathrm{y}}(s)}}, \quad K(\lambda) = -\frac{2\pi}{\sqrt{-T_{\mathrm{y}}(b)}},$$

with $T_{\rm y}(s) = (c-s)(\lambda-s)(a-s)$. Here, $\rho_{\rm y}(\lambda) = \rho(\lambda; c, a)$ is the rotation function of the ellipse obtained by sectioning the ellipsoid Q with the coordinate plane $\{y = 0\}$.

We begin with the limit $\epsilon := \lambda - b \to 0^+$. Using lemmas 15 and 18 we see that:

- $I(b+\epsilon) = I_0 + \mathrm{O}(\epsilon)$, where $I_0 = \int_0^c (b-s)^{-3/2}(c-s)^{-1/2}(a-s)^{-1/2}\mathrm{d}s$;
- $J(b+\epsilon) = 2\pi(a-b)^{-1/2}(b-c)^{-1/2}\epsilon^{-1/2} + \mathrm{O}(1)$;
- $K(b+\epsilon) = -2\pi(a-b)^{-1/2}(b-c)^{-1/2}\epsilon^{-1/2}$;
- $\rho_{\rm y}(b+\epsilon) = \rho_{\rm y}(b) + \mathrm{O}(\epsilon)$; and
- $\nu_{\rm y}(b+\epsilon) = \rho_{\rm z}(c) + \mathrm{O}(\epsilon^{1/2})$, since $\rho_{\rm y}(b) = \rho(b; c, a) = \rho(c; b, a) = \rho_{\rm z}(c)$.

Next, let us consider the case $\epsilon := a - \lambda \to 0^+$, which is similar to the limit $\lim_{\epsilon\to 0^+} \nu_{\rm x}(b-\epsilon)$ studied in the previous subsection, so we need the same simple integrals. Using them, jointly with lemmas 15 and 16, we get:

- $I(a-\epsilon) = I_* + \mathrm{O}(\epsilon)$, where $I_* = 2\pi(a-b)^{-1}\left((b-c)^{-1/2}\tilde{\rho}(c,b) - (a-c)^{-1/2}\tilde{\rho}(c,a)\right)$;
- $J(a-\epsilon) = J_* + \mathrm{O}(\epsilon)$, where $J_* = 2\pi(a-b)^{-1}(a-c)^{-1/2}$;
- $K(a-\epsilon) = K_* + \mathrm{O}(\epsilon)$, where $K_* = -2\pi(a-b)^{-1}(b-c)^{-1/2}$;
- $\rho_{\rm y}(a-\epsilon) = \tilde{\rho}(c,a) + \mathrm{O}(\epsilon)$; and
- $\nu_{\rm y}(a-\epsilon) = \tilde{\rho}(c,b) + \mathrm{O}(\epsilon) = \rho(b; c, b) + \mathrm{O}(\epsilon) = \rho_{\rm x}(b) + \mathrm{O}(\epsilon)$.

Appendix A.10. Asymptotic behaviour of the function $\nu_{\rm z}$

The function $\nu_{\rm z} : (0,c) \cup (c,b) \to \mathbb{R}$ verifies that

$$I(\lambda) + J(\lambda)\rho_{\rm z}(\underline{m}) + K(\lambda)\nu_{\rm z}(\lambda) = 0, \tag{A.11}$$

where the coefficients $I, J, K : (0,c) \cup (c,b) \to \mathbb{R}$ were given by

$$I(\lambda) = \int_0^{\overline{m}} \frac{\mathrm{d}s}{(\overline{m}-s)\sqrt{T_{\rm z}(s)}}, \quad J(\lambda) = 2\int_b^a \frac{\mathrm{d}s}{(s-\overline{m})\sqrt{T_{\rm z}(s)}}, \quad K(\lambda) = -\frac{2\pi}{\sqrt{-T_{\rm z}(\overline{m})}},$$

with $T_{\rm z}(s) = (\underline{m}-s)(b-s)(a-s)$, $\underline{m} = \min(\lambda, c)$, and $\overline{m} = \max(\lambda, c)$. Here, $\rho_{\rm z}(\lambda) = \rho(\lambda; b, a)$ is the rotation function of the ellipse obtained by sectioning the ellipsoid Q with the coordinate plane $\{z = 0\}$.

First, let us consider the case $\epsilon := \lambda \to 0^+$. Using lemmas 15 and 17 we see that:

- $I(\epsilon) = I_0\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2})$, where $I_0 = 2c^{-1}(ab)^{-1/2}$;
- $J(\epsilon) = J_0 + \mathrm{O}(\epsilon)$, where $J_0 = 2\int_b^a (s-c)^{-1}(s(s-b)(a-s))^{-1/2}\mathrm{d}s$;
- $K(\epsilon) = K_0 + \mathrm{O}(\epsilon)$, where $K_0 = -2\pi(c(b-c)(a-c))^{-1/2}$;
- $\rho_{\rm z}(\underline{m}) = \rho_{\rm z}(\min(\epsilon, c)) = \rho_{\rm z}(\epsilon) = \kappa^G\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2})$, where the constant $\kappa^G = \kappa^G(b,a)$ can be found in proposition 8; and
- $\nu_{\rm z}(\epsilon) = -(I_0 + J_0\kappa^G)K_0^{-1}\epsilon^{1/2} + \mathrm{O}(\epsilon^{3/2}) = \mathrm{O}(\epsilon^{1/2})$, with $(I_0 + J_0\kappa^G)K_0^{-1} < 0$.

The estimates in the limit $\epsilon := \overline{m} - \underline{m} \to 0^+$, which equivals to $\lambda \to c$, are:

- $I(c \pm \epsilon) = \pi(a-c)^{-1/2}(b-c)^{-1/2}\epsilon^{-1/2} + \mathrm{O}(1)$; see lemma 18;
- $J(c \pm \epsilon) = \mathrm{O}(1)$;

- $K(c \pm \epsilon) = -2\pi(a-c)^{-1/2}(b-c)^{-1/2}\epsilon^{-1/2} + \mathrm{O}(\epsilon^{1/2})$;
- $\rho_{\mathrm{z}}(\underline{m}) = \rho_{\mathrm{z}}(\min(c \pm \epsilon, c)) = \rho_{\mathrm{z}}(c) + \mathrm{O}(\epsilon)$, since $\rho_{\mathrm{z}}(\lambda)$ is analytic at $\lambda = c$; and
- $\nu_{\mathrm{z}}(c \pm \epsilon) = 1/2 + \mathrm{O}(\epsilon^{1/2})$.

Next, we consider the case $\epsilon := b - \lambda \to 0^+$. Using lemmas 15 and 16 we get:

- $I(b-\epsilon) = \mathrm{O}(1)$;
- $J(b-\epsilon) = J_*\epsilon^{-1/2} + \mathrm{O}(1)$, where $J_* = 2\pi(a-b)^{-1/2}(b-c)^{-1/2}$;
- $K(b-\epsilon) = K_*\epsilon^{-1/2} + \mathrm{O}(\epsilon^{1/2})$, where $K_* = -2\pi(a-b)^{-1/2}(b-c)^{-1/2}$;
- $\rho_{\mathrm{z}}(\underline{m}) = \rho_{\mathrm{z}}(\min(b-\epsilon, c)) = \rho_{\mathrm{z}}(c)$; and
- $\nu_{\mathrm{z}}(b-\epsilon) = \rho_{\mathrm{z}}(c) + \mathrm{O}(\epsilon^{1/2})$.

Appendix A.11. Asymptotic behaviour of the function $c_{\mathbf{0}}$

Let $\omega^0 = (\omega_1^0, \omega_2^0) \in \Omega$. We recall that the functions $\lambda = \lambda_{\mathbf{0}}(b)$ and $c = c_{\mathbf{0}}(b)$ are defined through the implicit equations

$$\nu_{\mathrm{z}}(\lambda) = \nu_{\mathrm{z}}(\lambda; c, b, a) = \omega_1^0, \qquad \rho_{\mathrm{z}}(\lambda) = \rho(\lambda; b, a) = \omega_2^0. \tag{A.12}$$

We want to prove that $\lim_{b\to 0^+} c_{\mathbf{0}}(b) = 0$, $\lim_{b\to 0^+} \lambda_{\mathbf{0}}(b) = 0$, $\lambda_* := \lim_{b\to a^-} \lambda_{\mathbf{0}}(b) = as_2^0$, and $c_* := \lim_{b\to a^-} c_{\mathbf{0}}(b) = as_2^0/s_1^0$, where $s_j^0 = \sin^2 \pi\omega_j^0$. The first two limits are trivial, because $0 < \lambda < c < b$. Let us consider the last two limits.

The limit $\tilde{\rho}(\beta, \alpha) = \lim_{\gamma\to\alpha^-} \rho(\gamma; \beta, \alpha) = \lim_{\gamma\to\alpha^-} \rho(\beta; \gamma, \alpha)$ verifies the equalities $\tan^2 \pi\tilde{\rho}(\beta, \alpha) = \beta/(\alpha - \beta)$ and $\sin^2 \pi\tilde{\rho}(\beta, \alpha) = \beta/\alpha$, see Appendix A.8. Hence, passing to the limit the second implicit equation in (A.12) we get that $\tilde{\rho}(\lambda_*, a) = \omega_2^0$, so $\lambda_* = as_2^0$. Next, by combining the two implicit equations in (A.12) with equality (A.11), we get

$$I(\lambda) + J(\lambda)\omega_2^0 + K(\lambda)\omega_1^0 = 0, \tag{A.13}$$

where $I(\lambda) = I(\lambda; c, b, a)$, $J(\lambda) = J(\lambda; c, b, a)$, and $K(\lambda) = K(\lambda; c, b, a)$ are quantities defined in Appendix A.10.

Using lemma 16 and some integrals computed in Appendix A.8, we have:

- $\lim_{b\to a} I(\lambda; c, b, a) = 2\pi(a-c)^{-1}\left((c-\lambda)^{-1/2}\tilde{\rho}(\lambda, c) - (a-\lambda)^{-1/2}\tilde{\rho}(\lambda, a)\right)$;
- $\lim_{b\to a} J(\lambda; c, b, a) = 2\pi(a-c)^{-1}(a-\lambda)^{-1/2}$; and
- $\lim_{b\to a} K(\lambda; c, b, a) = -2\pi(a-c)^{-1}(c-\lambda)^{-1/2}$.

Therefore, passing to the limit the identity (A.13) we get that

$$(c_* - \lambda_*)^{-1/2}\left(\tilde{\rho}(\lambda_*, c_*) - \omega_1^0\right) = (a - \lambda_*)^{-1/2}\left(\tilde{\rho}(\lambda_*, a) - \omega_2^0\right) = 0.$$

Thus $\rho(\lambda_*, c_*) = \omega_1^0$, which is equivalent to $\lambda_*/c_* = s_1^0$, so $c_* = \lambda_*/s_1^0 = as_2^0/s_1^0$.

Appendix B. A topological lemma

We recall that the complement of any Jordan curve X in the plane $\mathbb{R}^2$ has two distinct connected components. One of them is bounded and simply connected (the interior, denoted by $\mathcal{B}_X$) and the other is unbounded (the exterior, denoted by $\mathcal{U}_X$).

Lemma 22. *Let X and Y be two Jordan curves of $\mathbb{R}^2$. If $f : \mathcal{B}_X \to \mathbb{R}^2$ is a local homeomorphism whose image is bounded and that has a continuous extension to the boundary X such that $f(X) \subset Y$, then $f : \mathcal{B}_X \to \mathcal{B}_Y$ is a global homeomorphism.*

Proof. We note that $W = f(\mathcal{B}_X)$ is a non-empty open bounded subset of $\mathbb{R}^2$ such that

$$\partial W = \partial f(\mathcal{B}_X) \subset f(\partial \mathcal{B}_X) = f(X) \subset Y.$$

Next, we are going to prove that $W = \mathcal{B}_Y$. Using that $\partial W \subset Y$, we deduce that the intersection $W \cap \mathcal{B}_Y$ (respectively, $W \cap \mathcal{U}_Y$) is open and closed in $\mathcal{B}_Y$ (respectively, in $\mathcal{U}_Y$), so it is either the empty set or the whole interior (respectively, exterior). Therefore, we deduce that: 1) $W \cap \mathcal{U}_Y = \emptyset$, because W is bounded; 2) $W \cap Y = \emptyset$, because W is open; and 3) $W \cap \mathcal{B}_Y = \mathcal{B}_Y$, because W is open and non-empty. That is, $f(\mathcal{B}_X) = W = \mathcal{B}_Y$.

Once we know that $f : \mathcal{B}_X \to \mathcal{B}_Y$ is a surjective local homeomorphism, we deduce from covering space theory that it is a global homeomorphism. It suffices to realize that $\mathcal{B}_X$ is connected and open, and $\mathcal{B}_Y$ is simply connected. □

In particular, if $f : \mathcal{B}_X \to \mathbb{R}^2$ is smooth or analytic, then its inverse is also smooth or analytic. This means that if f is a local diffeomorphism whose image is bounded and that has a continuous extension to the boundary X such that $f(X) \subset Y$, then $f : \mathcal{B}_X \to \mathcal{B}_Y$ is a global diffeomorphism.

References

[1] S. Abenda and Yu. Fedorov, Closed geodesics and billiards on quadrics related to elliptic KdV solutions, *Lett. Math. Phys.*, **76**:111–134 (2006).

[2] M. Audin, Courbes algébriques et sistèmes intégrables: géodésiques des quadriques, *Expo. Math.*, **12**:193–226 (1994).

[3] M. Audin, Topologie des systèmes de Moser en dimension quatre, *The Floer Memorial Volume*, Progr. Math. vol. 133, 109–122 (1995).

[4] I. Babenko, Periodic trajectories in three-dimensional Birkhoff billiards, *Math. USSR-Sb.*, **71**:1–13 (1992).

[5] C. Batut, K. Belabas, D. Bernardi, H. Cohen and M. Olivier, *Users Guide to PARI/GP* (freely available from http://www.parigp-home.de/).

[6] M. Berger, Seules les quadriques admettent des caustiques, *Bull. Soc. Math. France*, **123**:107–116 (1995).

[7] G. D. Birkhoff, *Dynamical Systems*, Am. Math. Soc. Coll. Pub., vol. 9, 1927.

[8] S. Bolotin, A. Delshams and R. Ramírez-Ros, Persitence of homoclinic orbits for billiards and twist maps, *Nonlinearity*, **17**:1153–1177 (2004).

[9] P. S. Casas and R. Ramírez-Ros, Classification of symmetric periodic billiard trajectories inside ellipsoids, In preparation.

[10] S.-J. Chang and R. Friedberg, Elliptical billiards and Poncelet's theorem, *J. Math. Phys.*, **29**:1537–1550 (1988).

[11] S.-J- Chang, B. Crespi and K.-J. Shi, Elliptical billiard systems and the full Poncelet's theorem in n dimensions, *J. Math. Phys.*, **34**:2242–2256 (1993).

[12] B. Crespi, S.-J- Chang and K.-J. Shi, Elliptical billiards and hyperelliptic functions, *J. Math. Phys.*, **34**:2257–2289 (1993).
[13] A. Delshams, Yu. Fedorov and R. Ramírez-Ros, Homoclinic billiard orbits inside symmetrically perturbed ellipsoids, *Nonlinearity*, **14**:1141–1195 (2001).
[14] V. Dragović and M. Radnović, Conditions of Cayley's type for ellipsoidal billiard, *J. Math. Phys.*, **39**:355–362 (1998).
[15] V. Dragović and M. Radnović, On periodical trajectories of the billiard systems within an ellipsoid in $\mathbb{R}^d$ and generalized Cayley's condition, *J. Math. Phys.*, **39**:5866–5869 (1998).
[16] V. Dragović and M. Radnović, Geometry of integrable billiards and pencils of quadrics, *J. Math. Pures Appl.*, **85**:758–790 (2006).
[17] V. Dragović and M. Radnović, Hyperelliptic Jacobians as billiard algebra of pencils of quadrics: Beyond Poncelet porisms, *Adv. Math.*, **219**:1577–1607 (2008).
[18] V. Dragović and M. Radnović, Bifurcations of Liouville tori in elliptical billiards, *Regul. Chaotic Dyn.*, **14**:479–494 (2009).
[19] M. Farber, Topology of billiard problems (I, II), *Duke Math. J.*, **115**:559–585, 587–621 (2002).
[20] M. Farber and S. Tabachnikov, Periodic trajectories in 3-dimensional convex billiards *Manuscripta Math.*, **108**:431–437 (2002).
[21] Yu. Fedorov, Classical integrable systems and billiards related to generalized Jacobians, *Acta Appl. Math.*, **55**:251–301 (1999).
[22] Yu. Fedorov, Algebraic closed geodesics on a triaxial ellipsoid, *Regul. Chaotic Dyn.*, **10**:463–485 (2005).
[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1996.
[24] P. A. Griffiths, *Introduction to Algebraic Curves*, Transl. Math. Monographs, vol. 76, 1989.
[25] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.
[26] O. Knill, On nonconvex caustics of convex billiards, *Elem. Math.*, **53**:89–106 (1998).
[27] V. Kozlov V and D. Treshchëv, *Billiards: a Genetic Introduction to the Dynamics of Systems with Impacts*, Transl. Math. Monographs, vol. 89, 1991.
[28] H. Knörrer, Geodesics on the ellipsoid, *Inv. Math.*, **59**:119–143 (1980).
[29] H. Knörrer, Singular fibres of the momentum mapping for integrable Hamiltonian systems, *J. Mathematik*, **355**:67–107 (1985).
[30] R. Kolodziej, The rotation number of some transformation related to billiards in an ellipse, *Studia Math.*, **81**:293–302 (1985).
[31] J. K. Moser, Geometry of quadrics and spectral theory, *The Chern Symposium*, 147–188 (1980).
[32] G. Popov and P. Topalov, On the integral geometry of Liouville billiard tables, *Preprint* arXiv:0906.0451v1
[33] E. Previato, Poncelet theorem in space, *Proc. Amer. Math. Soc.*, **127**:2547–2556 (1999).
[34] R. Ramírez-Ros, Break-up of resonant invariant curves in billiards and dual billiards associated to perturbed circular tables, *Phys. D*, **214**:78–87 (2006).
[35] R. Ramírez-Ros, On Cayley-like conditions for elliptic billiards, In preparation.
[36] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems, *Regul. Chaotic Dyn.*, **6**:119–204 (2001).
[37] S. Tabachnikov, *Billiards*, Panoramas et Syntheses, vol. 1, Société Mathématique de France, 1995.
[38] M. B. Tabanov, Separatrices splitting for Birkhoffs billiard in symmetric convex domain, closed to an ellipse, *Chaos*, **4**:595–606 (1994)
[39] A. P. Veselov, Integrable systems with discrete time and difference operators, *Funct. Anal. Appl.*, **22**:83–93 (1988).
[40] A. P. Veselov, Integrable maps, *Russian Math. Surveys*, **46**:3–45 (1991).
[41] H. Waalkens, J. Wiersig and H. R. Dullin, The elliptic quantum billiard, *Ann. Phys. (NY)*, **260**:50–90 (1997).
[42] H. Waalkens, J. Wiersig and H. R. Dullin, Triaxial ellipsoidal quantum billiards, *Ann. Phys. (NY)*, **276**:64–110 (1999).
[43] H. Waalkens and H. R. Dullin, Quantum monodromy in prolate ellipsoidal billiards, *Ann. Phys. (NY)*, **295**:81–112 (2002).