Transfer operator approach to ray-tracing in circular domains

J Slipantschuk[1], M Richter[2,3], D J Chappell[3], G Tanner[2], W Just[1] and O F Bandtlow[1]

[1] School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK
[2] School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK
[3] School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, UK

E-mail: `j.slipantschuk@qmul.ac.uk`, `martin.richter@nottingham.ac.uk`, `david.chappell@ntu.ac.uk`, `gregor.tanner@nottingham.ac.uk`, `w.just@qmul.ac.uk`, `o.bandtlow@qmul.ac.uk`

Abstract. The computation of wave-energy distributions in the mid-to-high frequency regime can be reduced to ray-tracing calculations. Solving the ray-tracing problem in terms of an operator equation for the energy density leads to an inhomogeneous equation which involves a Perron-Frobenius operator defined on a suitable Sobolev space. Even for fairly simple geometries, let alone realistic scenarios such as typical boundary value problems in room acoustics or for mechanical vibrations, numerical approximations are necessary. Here we study the convergence of approximation schemes by rigorous methods. For circular billiards we prove that convergence of finite-rank approximations using a Fourier basis follows a power law where the power depends on the smoothness of the source distribution driving the system. The relevance of our studies for more general geometries is illustrated by numerical examples.

AMS classification scheme numbers: 37M25, 37C30, 74H45, 37D50

1. Introduction

Ray-tracing methods serve as an important toolkit in finding approximate solutions of linear wave equations in the high frequency limit. This approximation is used in a variety of fields providing, for example, the connection between Maxwell's equations and geometric optics, as well as between quantum mechanics and classical Hamiltonian mechanics [16]. The ray-tracing limit has also been considered in detail in acoustics, seismology and mechanical vibrations [27]. In engineering applications, ray tracing is employed in handling electromagnetic problems, such as coverage estimates for 5G or WiFi communication [14], room acoustics simulations [26] as well as structure-borne sound propagation in mechanical structures [8]. Finding closed form, analytical solutions to such engineering problems of sufficient complexity is generally impossible, even using ray-tracing techniques, and one has to use numerical methods instead.

For solving linear wave problems such as those listed above, the numerical methods used have to be adapted to the relevant length and frequency scales involved. In the low frequency regime, finite element methods (FEM) are routinely employed for resolving the full wave dynamics. However, the number of degrees of freedom in an FEM model needs to scale with the wavelength and there is thus an upper limit in frequency above which the required computational resources become unfeasible. At very high frequencies, power balance approaches can often be used as long as certain assumptions on the ergodicity of the underlying ray dynamics are satisfied [28]. In the mid-to-high frequency range, ray-tracing becomes the method of choice; standard ray-tracing techniques track all possible rays from a source to a receiver point [26] — a method which becomes cumbersome if many reflections need to be taken into account. As an alternative *Dynamical Energy Analysis* (DEA) was proposed and has proven to be useful in particular for structure-borne sound problems [28, 17]. Instead of tracking individual rays carrying vibrations across the complex structure — which is extremely challenging — in DEA, the problem is reformulated in terms of densities of rays, which are then mapped across a mesh representing the structure [9, 10]. This reduces the ray tracing problem from tracking rays on complicated and curved domains to mapping ray segments across small, plane patches of a simple shape forming the mesh, typically triangular or quadrilateral mesh cells. The ray densities are then mapped from one cell of the mesh to adjacent ones and the overall transport problem can be formulated in terms of an inhomogeneous equation of the form

$$(I - \mathcal{L})f = f_0 \,, \tag{1}$$

where f_0 is the initial ray density, $\mathcal{L}$ a Perron-Frobenius type operator describing the evolution of ray densities and f the required final ray density. Using DEA, the distribution of vibrational energy in mechanical structures, such as ships, cars and tractors [18, 17] can be calculated successfully.

For such realistic geometries, Equation (1) above cannot be solved analytically, so recourse is made to numerical schemes based on heuristic finite-dimensional matrix approximations of the operator $\mathcal{L}$. To date, very little is known about the convergence

properties of these schemes and the dependence of the convergence rate on the ray dynamics, as well as the discretisation techniques [10]. The precise form of convergence is likely to be highly sensitive to both the basis functions used in approximating the inhomogeneous Equation (1), as well as dynamical and damping properties of the system under investigation [18]. For our study, we will therefore be concerned with the approximation of $\mathcal{L}$ by operators of finite rank. There is a plethora of papers on numerical approximation of Perron-Frobenius operators, starting with Ulam's method of phase space discretisation, finite section or Galerkin methods, and data-driven methods, see for example [2, 12, 13, 20, 22] to mention but a few. Surprisingly, the application of DEA (which falls into the Galerkin category) to even fairly simple geometries has not been dealt with at a rigorous level. Here, we shall thus focus on one of the simplest cases, the billiard dynamics given by the ballistic motion within a circular disk. We shall establish rigorous error bounds of finite-dimensional approximations for the resulting energy distribution.

In order to set up the required notation, consider a particle moving inside a circular billiard table $\mathcal{D}$ being specularly reflected at its boundary $\partial\mathcal{D}$. We parametrise $\partial\mathcal{D}$ by the polar angle $x \in \mathbb{R}/2\pi\mathbb{Z}$ and we denote by $y \in [-\pi/2, \pi/2]$ the angle of reflection that the postcollisional velocity vector has with the inward normal to $\partial\mathcal{D}$. Initially the collision angle is defined on an interval. It is, however, technically simpler to deal with cyclic variables. Since both angles $-\pi/2$ and $\pi/2$ correspond to a particle which sticks on the boundary we identify both angles so that the collision angle becomes a cyclic variable as well. With these conventions, the collision map T on the domain $\Omega = (\mathbb{R}/2\pi\mathbb{Z}) \times (\mathbb{R}/\pi\mathbb{Z})$ can be written as

$$T(x,y) = (x + \pi - 2y, y), \quad (x,y) \in \Omega \tag{2}$$

with its inverse $\phi = T^{-1}$ given by

$$\phi(x,y) = (x - \pi + 2y, y) \quad (x,y) \in \Omega. \tag{3}$$

It is not difficult to see that the collision map T preserves the normalised Lebesgue measure on Ω. The long-term statistical behaviour of T can thus be studied by investigating the associated Perron-Frobenius operator (see, for example, [7]), which for invertible measure-preserving maps is given by the composition operator $\mathcal{C}_\phi$ defined as

$$(\mathcal{C}_\phi f)(x,y) = f(\phi(x,y)), \quad (x,y) \in \Omega, \tag{4}$$

where $f \colon \Omega \to \mathbb{C}$. In the current work we are interested in the properties of a weighted Perron-Frobenius operator, also known as a transfer operator. In order to define it, let us first introduce a multiplication operator $\mathcal{M}_w$ acting on functions $f \colon \Omega \to \mathbb{C}$ by

$$(\mathcal{M}_w f)(x,y) = w(x,y)f(x,y), \quad (x,y) \in \Omega, \tag{5}$$

where $w \colon \Omega \to [0, \infty)$ is a suitable weight function, which in the DEA framework accounts for dissipation caused either by collisions with the wall or by in-flight

dissipation. The transfer operator, understood to be acting on a suitable space of functions detailed in the following section, is now given by

$$\mathcal{L} = \mathcal{M}_w \mathcal{C}_\phi \,. \tag{6}$$

In the present article, we are interested in approximations of the solution to the operator Equation (1) with $f_0 \colon \Omega \to [0, \infty)$ interpreted as the initial boundary density of particles induced by the first boundary collision of particles emitted by a source located in the interior of $\mathcal{D}$ (see [28]). In the DEA approach this quantity represents the energy source. The resulting energy distribution is captured by the solution, $f \colon \Omega \to [0, \infty)$, which gives the stationary boundary density generated by the collision dynamics. Given a suitable Banach space and a sequence of finite-rank projections $(P_K)_{K \in \mathbb{N}}$, an approximation method for (1) can be constructed by considering the projected finite-dimensional problem

$$(I - P_K \mathcal{L} P_K) f_K = f_0 \,. \tag{7}$$

The aim of this work is to present a Banach space for f_0 and $(P_K)_{K \in \mathbb{N}}$, so that problem (7) has solutions, which converge in a suitable topology to the solution of (1) as K tends to infinity, with the speed of convergence being of the order $K^{-\alpha}$. The exponent α depends on the smoothness of f_0 and the requirements imposed on the type of convergence.

In passing we note that transfer operators have their roots in statistical mechanics [23, 24] and nowadays play an important role in the ergodic theory of smooth expanding, or more generally, hyperbolic dynamical systems (see, for example, [3, 4]). The main reason for their popularity in this context derives from the fact that for expanding or hyperbolic dynamical systems the transfer operator, when considered on a suitable function space, can be shown to have discrete peripheral spectrum, from which long-term statistical properties of the underlying system can be derived. In the elliptic setting, however, such as for the circular billiard considered in this article, analogous results cannot be expected, and, as a consequence, transfer operator methods have received little attention in this context. It is perhaps worth noting that in our setting we do not require discreteness of the peripheral spectrum of the transfer operator. The main onus is to show that the resolvent of the transfer operator exists at the point 1 (see Equation (1)) and can be effectively approximated by finite-rank operators (see Equation (7)).

As we intend to keep our presentation accessible to non-specialists, we will occasionally elaborate on aspects covered in the specialised literature but which may not be well known to a general audience. The remaining parts are organised as follows. In Section 2 we introduce Sobolev spaces, on which the transfer operator and its finite-dimensional approximations are bounded operators with spectral radii bounded away from 1. In Section 3 we shall prove the convergence results for the operator equations (1) and (7) stated as Theorem 3.4. In the final Section 4 we summarise the main findings, compare the formal results with numerical simulations and explore the relevance of the current study in a wider context.

2. Sobolev spaces and transfer operators

We will be interested in certain subspaces of $L^2(\Omega) = L^2(\Omega, m)$ where $dm = dxdy/(2\pi^2)$ is the normalised two-dimensional Lebesgue measure on Ω. The natural inner product is given by

$$(f, g)_{L^2} = \int_\Omega f(x, y)\overline{g(x, y)}dm.$$

An orthonormal basis of $L^2(\Omega)$ is given by $\{e_k \colon k \in \mathbb{Z}^2\}$ where $e_k(x, y) = \mathrm{e}^{ik_1 x}\mathrm{e}^{2ik_2 y}$ so that $f(x, y) = \sum_{k \in \mathbb{Z}^2} c_k(f)e_k(x, y)$ with Fourier coefficients $c_k(f) = (f, e_k)_{L_2}$.

Definition 2.1. Let $m = (m_1, m_2) \in \mathbb{N}_0^2$. The Sobolev space $H^m(\Omega)$ is the collection of all $f \in L^2(\Omega)$ such that for all $\nu = (\nu_1, \nu_2) \in \mathbb{N}^2$ with $\nu_1 \le m_1$ and $\nu_2 \le m_2$ the weak derivatives $D^\nu f = D_x^{\nu_1} D_y^{\nu_2} f$ exist and belong to $L^2(\Omega)$.

The space $H^m(\Omega)$ is a Hilbert space, when equipped with the inner product‡

$$(f, g)_{H^m} = (f, g)_{L^2} + (D_x^{m_1} f, D_x^{m_1} g)_{L^2} + (D_y^{m_2} f, D_y^{m_2} g)_{L^2}. \tag{8}$$

One can rewrite this definition in terms of Fourier coefficients. Using the fact that $c_k(D^\nu f) = (ik_1)^{\nu_1}(2ik_2)^{\nu_2}c_k(f)$, Equation (8) can be expressed as

$$(f, g)_{H^m} = \sum_{k \in \mathbb{Z}^2}(1 + |k_1|^{2m_1} + |2k_2|^{2m_2})c_k(f)\overline{c_k(g)}. \tag{9}$$

Remark 2.2. For $m = (m_1, m_2)$ with $m_1 = m_2$ the Sobolev space $H^m(\Omega)$ coincides with the classical isotropic Sobolev space, while for $m_1 \ne m_2$, the space is an example of an anisotropic Sobolev space (see, for example, [11, Sec. 2.2]).

Using Equation (9) we can define fractional Sobolev spaces $H^s(\Omega)$ for $s = (s_1, s_2) \in \mathbb{R}_+^2$ as

$$H^s(\Omega) = \left\{ f \in L^2(\Omega) \colon \sum_{k \in \mathbb{Z}^2} |c_k(f)|^2(1 + |k_1|^{2s_1} + |2k_2|^{2s_2}) < \infty \right\},$$

which are Hilbert spaces when equipped with the inner product given in Equation (9) with m replaced by s.

We shall next investigate the properties of the composition operator $\mathcal{C}_\phi$ associated with the map ϕ in (3) on the fractional Sobolev space $H^s(\Omega)$.

Lemma 2.3. *The composition operator $\mathcal{C}_\phi$ given in (4) considered on $H^s(\Omega)$ with $s_1 \ge s_2 \ge 0$ is bounded, with spectral radius $r(\mathcal{C}_\phi) = 1$.*

Proof. For any $n \in \mathbb{N}$ and $(x, y) \in \Omega$ we have $\phi^n(x, y) = (x - n\pi + 2ny, y)$, and thus

$$(\mathcal{C}_\phi^n e_k)(x, y) = (\mathcal{C}_{\phi^n} e_k)(x, y) = (-1)^{k_1 n} e_{k_1, nk_1 + k_2}(x, y), \tag{10}$$

for any $k \in \mathbb{Z}^2$.

‡ This choice of inner product is sometimes referred to as the modified inner product, in contrast with the classical one (see, for example, [21, Def 2.2]).

In order to show that the operator is bounded we will need the following general inequality. Let $(x, y) \in [0, \infty)^2$ and $t \geq 0$, then

$$(x + y)^t \leq C_t(x^t + y^t), \quad \text{with } C_t = \max(1, 2^{t-1}). \tag{11}$$

Using Equation (11) we obtain the bound $|nk_1 + k_2|^{2s_2} \leq C_{2s_2}(n^{2s_2}|k_1|^{2s_1} + |k_2|^{2s_2})$ for $s_1 \geq s_2$, which leads to

$$\|\mathcal{C}_\phi^n e_k\|_{H^s}^2 = 1 + |k_1|^{2s_1} + |2(nk_1 + k_2)|^{2s_2} \leq \left(1 + C_{2s_2}(2n)^{2s_2}\right) \|e_k\|_{H^s}^2.$$

Since $(\mathcal{C}_\phi^n e_k, \mathcal{C}_\phi^n e_l)_{H^s} = 0$ for $k \neq l$, the operator norm of $\mathcal{C}_\phi^n$ is bounded from above by $(1 + (2n)^{2s_2} \max(1, 2^{2s_2-1}))^{1/2}$, resulting in the upper bound for the spectral radius

$$r(\mathcal{C}_\phi) = \lim_{n \to \infty} \|\mathcal{C}_\phi^n\|^{1/n} \leq \lim_{n \to \infty} \left(1 + (2n)^{2s_2} \max(1, 2^{2s_2-1})\right)^{1/(2n)} = 1.$$

In order to see that the inequality above is an equality, observe that the operator norm of $\mathcal{C}_\phi^n$ is bounded from below by 1 as $\|\mathcal{C}_\phi^n e_0\|_{H^s} = \|e_0\|_{H^2}$. Thus $r(\mathcal{C}_\phi) = 1$. □

Before proceeding we note that by (10), the action of the composition operator on $H^s(\Omega)$ can be represented by the action of the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

on Fourier coefficients. In particular, we have

$$\mathcal{C}_\phi^n e_k = (-1)^{k_1 n} e_{A^n k}. \tag{12}$$

For $K \in \mathbb{N}$ define $\Lambda_K = \Lambda_K^0 = \{(k_1, k_2) \in \mathbb{Z}^2 : |k_1| < K, |k_2| < K\}$, and let $\Lambda_K^n = A^n(\Lambda_K)$. Then for any $n \in \mathbb{N}_0$ we can define a finite-rank operator $P_{\Lambda_K^n} : H^s(\Omega) \to H^s(\Omega)$ by

$$(P_{\Lambda_K^n} f)(x, y) = \sum_{k \in \Lambda_K^n} c_k(f) e_k(x, y), \quad (x, y) \in \Omega. \tag{13}$$

Lemma 2.4. *Let $\mathcal{C}_\phi$ and P_{Λ_K} be as above. Then*

$$\mathcal{C}_\phi^n P_{\Lambda_K} = P_{\Lambda_K^n} \mathcal{C}_\phi^n$$

for any $n, K \in \mathbb{N}_0$.

Proof. This follows by checking the equality for any basis element e_k and noting that A^n is invertible. □

Definition 2.5. Let $\mathcal{M}_w$ denote the multiplication operator as defined in Equation (5), considered as an operator on $H^s(\Omega)$, with a smooth weight function $w \colon \Omega \to [0, \infty)$. In addition, we assume that w has the following properties:

(a) $\|w\|_\infty = \sup_{x \in \Omega} |w(x)| < 1$;

(b) w is bounded away from zero;

(c) $w(x, y) = w(x', y)$ for any $(x, y), (x', y) \in \Omega$, that is, the weight w does not depend on the first argument.

Remark 2.6. The operator $\mathcal{M}_w$ models the effect of damping on the motion of the billiard particle. Assumptions (a) and (b) imply that the damping is well-behaved, while assumption (c) is innocuous, given the circular symmetry of the billiard table.

The following two lemmas summarise basic properties of $\mathcal{M}_w$ and $\mathcal{C}_\phi$.

Lemma 2.7. *Let $\mathcal{M}_w, \mathcal{C}_\phi$ and P_{Λ_K} be as above. Then we have the following.*

(i) $\mathcal{M}_w \mathcal{C}_\phi = \mathcal{C}_\phi \mathcal{M}_w$;

(ii) $D_x \mathcal{C}_\phi = \mathcal{C}_\phi D_x$;

(iii) $D_x \mathcal{M}_w = \mathcal{M}_w D_x$;

(iv) $D_y \mathcal{C}_\phi^n = 2n \mathcal{C}_\phi^n D_x + \mathcal{C}_\phi^n D_y$ *for* $n \in \mathbb{N}$;

(v) $D_y \mathcal{M}_w^n = n \mathcal{M}_w^{n-1} \mathcal{M}_{D_y w} + \mathcal{M}_w^n D_y$ *for* $n \in \mathbb{N}$;

(vi) $D_x P_{\Lambda_K} = P_{\Lambda_K} D_x$ *and* $D_y P_{\Lambda_K} = P_{\Lambda_K} D_y$ *for* $K \in \mathbb{N}$.

Proof. Items *(i)* and *(iii)* follow from Definition 2.5(c); items *(ii)* and *(iv)* follow by direct computation using the map ϕ; item *(v)* is obvious and *(vi)* is a direct consequence of the relations $c_k(D_x f) = (\mathrm{i}k_1) c_k(f)$ and $c_k(D_y f) = (2\mathrm{i}k_2) c_k(f)$. $\qquad\square$

We write $\mathcal{L}_K = P_{\Lambda_K} \mathcal{M}_w \mathcal{C}_\phi P_{\Lambda_K}$ for the finite-rank approximation of $\mathcal{L} = \mathcal{M}_w \mathcal{C}_\phi$. Using Lemma 2.7 *(i)* and Lemma 2.4, we can write $\mathcal{L}_K^n$ for $n \in \mathbb{N}$ as

$$\mathcal{L}_K^n = (P_{\Lambda_K} \mathcal{M}_w \mathcal{C}_\phi P_{\Lambda_K})^n = P_{\Lambda_K} \left(\prod_{l=1}^n \mathcal{M}_w P_{\Lambda_K^l} \right) \mathcal{C}_\phi^n. \tag{14}$$

In order to state the properties of $\mathcal{L}$ and $\mathcal{L}_K$ we need to introduce the following multi-index notation: an n-dimensional multi-index is an n-tuple $\boldsymbol{i}_n = (i_1, i_2, \ldots, i_n)$ of non-negative integers of order $|\boldsymbol{i}_n| = i_1 + i_2 + \cdots + i_n = m$; the corresponding multinomial coefficient is given by

$$\binom{m}{\boldsymbol{i}_n} = \frac{m!}{i_1! i_2! \cdots i_n!}.$$

Lemma 2.8. *Let $\mathcal{M}_w, \mathcal{C}_\phi$ and $P_{\Lambda_K^l}$ be as above. Then we have the following.*

(i) $D_y^m \mathcal{C}_\phi = \sum_{i=0}^m 2^{m-i} \binom{m}{i} \mathcal{C}_\phi D_x^{m-i} D_y^i$;

(ii) $D_y^m \mathcal{C}_\phi^n = \sum_{|\boldsymbol{i}_{n+1}|=m} 2^{m-i_{n+1}} \binom{m}{\boldsymbol{i}_{n+1}} \mathcal{C}_\phi^n D_x^{m-i_{n+1}} D_y^{i_{n+1}}$;

(iii) $D_y^m \left(\prod_{l=1}^n \mathcal{M}_w P_{\Lambda_K^l} \right) = \sum_{|\boldsymbol{i}_{n+1}|=m} \binom{m}{\boldsymbol{i}_{n+1}} \left(\prod_{l=1}^n \mathcal{M}_{D_y^{i_l} w} P_{\Lambda_K^l} \right) D_y^{i_{n+1}}$.

Proof. Item *(i)* follows by induction over m using Lemma 2.7*(iv)* for the base case $m = 1$. For item *(ii)*, the additional induction over n follows by rewriting *(i)* as $D_y^m \mathcal{C}_\phi = \sum_{i_1+i_2=m} 2^{i_1} \binom{m}{i_1,i_2} \mathcal{C}_\phi D_x^{i_1} D_y^{i_2}$. Finally, item *(iii)* follows from the Leibniz formula. $\qquad\square$

We are now ready to prove the main result of this section. Keeping in mind that we assume that the billiard dynamics is dissipative, that is, the weight is chosen so that $\|w\|_\infty < 1$, the following lemma shows that, given $f_0 \in H^s(\Omega)$, the problem (1) and the projected version (7) have unique solutions $f \in H^s(\Omega)$ and $f_K \in H^s(\Omega)$, respectively.

Lemma 2.9. *Consider $\mathcal{L}$ and $\mathcal{L}_K$, $K \in \mathbb{N}$, as operators on $H^s(\Omega)$ for $s \in \mathbb{N}_0^2$ with $s_1 \geq s_2 \geq 0$. Then*

(i) $(\mathcal{L}_K)_{K \in \mathbb{N}}$ is a family of bounded operators $H^s(\Omega)$ with norms bounded uniformly in K. Moreover, $r(\mathcal{L}_K) \leq \|w\|_\infty$ for all K;

(ii) $\mathcal{L}$ is a bounded operator on $H^s(\Omega)$ with $r(\mathcal{L}) \leq \|w\|_\infty$.

Proof. We shall only prove statement *(i)*, as the proof of statement *(ii)* follows by almost identical arguments. In the following, we shall assume that $s_1 \geq s_2 \geq 1$, as the case $s_1 s_2 = 0$ follows by identical arguments. For $f \in H^s(\Omega)$ we have

$$\|\mathcal{L}_K^n f\|_{H^s}^2 = \|\mathcal{L}_K^n f\|_{L^2}^2 + \|D_x^{s_1} \mathcal{L}_K^n f\|_{L^2}^2 + \|D_y^{s_2} \mathcal{L}_K^n f\|_{L^2}^2. \tag{15}$$

Let $p, q \in \mathbb{N}$ with $p \leq s_1$ and $q \leq s_2$. It is not difficult to see that for any $f \in H^s(\Omega)$ and $K \in \mathbb{N}_0$ the following holds.

(a) $\|P_{\Lambda_K^j} f\|_{L^2} \leq \|f\|_{L^2}$ for any $j \in \mathbb{N}_0$;

(b) $\|\mathcal{M}_w f\|_{L^2} \leq \|w\|_\infty \|f\|_{L^2}$;

(c) $\|D_x^p f\|_{L^2}^2 \leq \|D_x^{s_1} f\|_{L^2}^2$ and $\|D_x^q f\|_{L^2}^2 \leq \|D_x^{s_2} f\|_{L^2}^2$;

(d) $\|D_x^p D_y^q f\|_{L^2}^2 \leq \|D_x^{p+q} f\|_{L^2}^2 + \|D_y^{p+q} f\|_{L^2}^2$ wherever $p + q \leq s_2$.

Here, statements (c) and (d) follow by writing the L^2 norm of $D_x^p f$ and $D_y^q f$ using Parseval's identity.

Writing $\mathcal{L}_K^n$ as in Equation (14) and using (a) and (b) above iteratively we have

$$\|\mathcal{L}_K^n f\|_{L^2} = \|(P_{\Lambda_K} \mathcal{M}_w \mathcal{C}_\phi P_{\Lambda_K})^n f\|_{L^2} \leq \|w\|_\infty^n \|\mathcal{C}_\phi^n f\|_{L^2} \leq \|w\|_\infty^n \|f\|_{L^2}, \tag{16}$$

where the last inequality follows from the fact that the operator norm of $\mathcal{C}_\phi$ on $L^2(\Omega)$ equals 1.

As D_x commutes with any of the operators involved (Lemma 2.7 *(ii,iii,vi)*) we have in the second term on the right-hand side of (15) that $D_x^{s_1} \mathcal{L}_K^n = \mathcal{L}_K^n D_x^{s_1}$. By the same argument as above we have

$$\|D_x^{s_1} \mathcal{L}_K^n f\|_{L^2} \leq \|w\|_\infty^n \|D_x^{s_1} f\|_{L^2}. \tag{17}$$

In order to bound the last term in Equation (15) we are using Lemma 2.8*(iii)* and Hölder's inequality in order to write

$$\|D_y^{s_2} \mathcal{L}_K^n f\|_{L^2}^2 = \|\sum_{j=0}^{s_2} A_j D_y^j \mathcal{C}_\phi^n f\|_{L^2}^2 \leq (s_2 + 1) \sum_{j=0}^{s_2} \|A_j\|_{L^2}^2 \|D_y^j \mathcal{C}_\phi^n f\|_{L^2}^2,$$

where $A_j = \sum_{|\boldsymbol{i}_n|=s_2-j} \binom{s_2}{\boldsymbol{i}_n,j} \left(\prod_{l=1}^n M_{D_y^{i_l} w} P_{\Lambda_K^l} \right)$.

We shall first obtain a bound for $\|D_y^j \mathcal{C}_\phi^n f\|_{L^2}$. Using Lemma 2.8 *(ii)* and decomposing the sum in terms of powers of D_x and D_y we obtain

$$D_y^j \mathcal{C}_\phi^n = (2n)^j C_\phi^n D_x^j + C_\phi^n D_y^j + \sum_{\substack{|i_{n+1}|=j \\ 0<i_{n+1}<j}} 2^{j-i_{n+1}} \binom{j}{i_{n+1}} \mathcal{C}_\phi^n D_x^{j-i_{n+1}} D_y^{i_{n+1}},$$

where we have used the multinomial formula $\sum_{|i_n|=k} \binom{k}{i_n} = n^k$. Thus, for $j \le m$ we obtain using Hölder's inequality, the multinomial formula and upper bounds for $2^{j-i_{n+1}}$

$$\|D_y^j \mathcal{C}_\phi^n f\|_{L^2}^2 \le 2^j (n+1)^j \left(2^j n^j \|D_x^j f\|_{L^2}^2 + \|D_y^j f\|_{L^2}^2 \right)$$
$$+ 2^j (n+1)^j \left((2^j(n+1)^j - 2^j n^j - 1) \max_{0<i<j} \|D_x^{j-i} D_y^i f\|_{L^2}^2 \right)$$
$$\le 2^{2s_2}(n+1)^{2s_2} \left(\|D_x^{s_1} f\|_{L^2}^2 + \|D_y^{s_2} f\|_{L^2}^2 \right), \tag{18}$$

where the last inequality uses (c) and (d).

Next we shall obtain a bound on the operator norm of A_j for $j \le s_2$. First note that $\mathcal{M}_{D_y^l w} = \mathcal{M}_w \mathcal{M}_{(D_y^l w)/w}$ is well-defined as w is bounded away from zero. By using (a) and (b) iteratively, for any $\boldsymbol{i}_n = (i_1, \ldots, i_n)$ with $|\boldsymbol{i}_n| = s_2$ we have

$$\left\| \prod_{l=1}^n \mathcal{M}_{D_y^{i_l} w} P_{\Lambda_K^l} f \right\|_{L^2}^2 \le C_{s_2} \|w\|_\infty^{2n} \|f\|_{L^2}^2$$

where $C_{s_2} = \max_{\substack{i_1,\ldots,i_n \\ |\boldsymbol{i}_n|=s_2}} \left(\prod_{l=1}^n \|D_y^{i_l} w/w\|_\infty^2 \right) \le \max_{0 \le l \le s_2} \|D_y^l w/w\|_\infty^{2s_2}$ is a constant independent of n. Using arguments analogous to those used to obtain inequality (18), we obtain the bound

$$\|A_j\|_{L^2}^2 \le (n+1)^{2s_2} C_{s_2} \|w\|_\infty^{2n}. \tag{19}$$

Using the estimates (16), (17), (18) and (19) in Equation (15) we arrive at the bound

$$\|\mathcal{L}_K^n f\|_{H^s}^2 \le \tilde{C}_{n,s_2} \|w\|_\infty^{2n} \|f\|_{H^2}^2$$

with $\tilde{C}_{n,s_2} \le (s_2+1)s_2(n+1)^{4s_2} 2^{2s_2} C_{s_2} + 1$. As $\tilde{C}_{n,s_2}$ is independent of K, the family $(\mathcal{L}_K)_{K \in \mathbb{N}}$ is a uniformly bounded family of bounded operators on $H^s(\Omega)$. Finally, taking the right hand side of Equation (15) to the power of $1/n$ and observing that $\tilde{C}_{n,s_2}$ grows polynomially in n, the upper bound for the spectral radius of $\mathcal{L}_K$ follows. $\qquad\square$

3. Convergence properties

In the previous section we established (see Lemma 2.9) that given $f_0 \in H^s(\Omega)$, the problem (1) and the projected version (7) have unique solutions $f \in H^s(\Omega)$ and $f_K \in H^s(\Omega)$, respectively. We shall now turn to establishing the convergence of f_K to f. This would be straightforward if we knew that $\mathcal{L}_K \to \mathcal{L}$ as $K \to \infty$ in the operator norm on $H^s(\Omega)$, since then, using the so-called second resolvent identity

$$(I - \mathcal{L}_K)^{-1} - (I - \mathcal{L})^{-1} = (I - \mathcal{L}_K)^{-1}(\mathcal{L}_K - \mathcal{L})(I - \mathcal{L})^{-1}, \tag{20}$$

we would have

$$\|f_K - f\|_{H^s} = \|(I - \mathcal{L}_K)^{-1}f_0 - (I - \mathcal{L})^{-1}f_0\|_{H^s} = \|(I - \mathcal{L}_K)^{-1}(\mathcal{L}_K - \mathcal{L})(I - \mathcal{L})^{-1}f_0\|_{H^s},$$

from which convergence of $f_K \to f$ in $H^s(\Omega)$ could be readily obtained.

This, however, cannot be the case, as if $\mathcal{L}_K \to \mathcal{L}$ as $K \to \infty$ in the operator norm on $H^s(\Omega)$, then $\mathcal{L}$, as a uniform limit of finite-rank operators, would be compact on $H^s(\Omega)$. However, as $\mathcal{L}$ has a bounded inverse on $H^s(\Omega)$, it cannot be compact.

We thus need to resort to a slightly weaker notion of convergence, that is, we shall consider the transfer operator as an operator between Sobolev spaces of different order. In passing, we remark that this idea is also at the heart of one of the most successful techniques to obtain spectral approximation results of transfer operators, where perturbation sizes are measured in 'triple' norms (see, for example, [19]).

In the following we shall explain this idea in more detail. We start with the following important observation. For $t, s \in [0, \infty)^2$ with $s_1 \geq s_2 > t_1 \geq t_2 \geq 0$, functions in $H^s(\Omega)$ can be identified with functions in $H^t(\Omega)$ using the embedding operator $\mathcal{J} \colon H^s(\Omega) \hookrightarrow H^t(\Omega)$ given by $\mathcal{J}f = f$. This operator is not just continuous, but also compact, as the following lemma shows.

Lemma 3.1. *Let $\mathcal{J} \colon H^s(\Omega) \hookrightarrow H^t(\Omega)$ be the canonical embedding, where $t, s \in [0, \infty)^2$ with $s_1 \geq s_2 > t_1 \geq t_2 \geq 0$. Let $P_K = P_{\Lambda_K}$ the projection operator in Equation (13), and $\mathcal{J}_K = \mathcal{J}P_K$. Then,*

$$\|\mathcal{J} - \mathcal{J}_K\|_{H^s \to H^t} \leq C(1 + K^2)^{-\alpha/2}$$

for some $C > 0$ and $\alpha = (s_2 - t_1)$.

Proof. Let $f \in H^s(\Omega)$. Using the notation $a_t(k) = 1 + |k_1|^{2t_1} + |2k_2|^{2t_2}$ we have

$$\|\mathcal{J}f - \mathcal{J}_K f\|_{H^t}^2 = \sum_{i=1}^{3} \sum_{k \in I_i(K)} |c_k(f)|^2 a_t(k),$$

with $I_1(K) = \{k \in \mathbb{Z}^2 \colon |k_1| \geq K, |k_2| \geq K\}$, $I_2(K) = \{k \in \mathbb{Z}^2 \colon |k_1| < K, |k_2| \geq K\}$, $I_3(K) = \{k \in \mathbb{Z}^2 \colon |k_1| \geq K, |k_2| < K\}$. We will first show that there exists a constant C' such that

$$a_t(k) \leq C'(1 + |k_1|^2 + |2k_2|^2)^{-\alpha} a_s(k).$$

For this, first observe that

$$(1 + |k_1|^2 + |2k_2|^2)^{s_2} \leq C_{s_2}(1 + |k_1|^{2s_1} + |2s_2|^{2s_2}) \leq C_{s_2} a_s(k),$$

which follows by Hölder's inequality and $s_1 \geq s_2$. Then,

$$\begin{aligned}
a_t(k) = 1 + |k_1|^{2t_1} + |2k_2|^{2t_1} &\leq 3(1 + |k_1|^2 + |2k_2|^2)^{t_1} \\
&= 3(1 + |k_1|^2 + |2k_2|^2)^{t_1 - s_2}(1 + |k_1|^2 + |2k_2|^2)^{s_2} \\
&\leq 3C_{s_2}(1 + |k_1|^2 + |2k_2|^2)^{t_1 - s_2} a_s(k).
\end{aligned}$$

Now, by bounding from above each $(1 + |k_1|^2 + |2k_2|^2)^{-\alpha}$ with its maximal value in each of the sums, we obtain

$$\|\mathcal{J}f - \mathcal{J}_K f\|_{H^t}^2 \leq C' \left((1 + 5K^2)^{-\alpha} + (1 + 4K^2)^{-\alpha} + (1 + K^2)^{-\alpha} \right) \|f\|_{H^s}^2$$
$$\leq 3C'(1 + K^2)^{-\alpha} \|f\|_{H^s}^2. \qquad \square$$

We are now able to show that $\mathcal{L}$ can be approximated by finite-rank operators when considered as operators from H^s to H^t.

Proposition 3.2. *Let $\mathcal{L}_K = P_K \mathcal{L} P_K$ be the finite-rank approximation of $\mathcal{L}$ on $H^s(\Omega)$ with $s \in \mathbb{N}^2$ and $s_1 \geq s_2$. Let $\mathcal{J}$ be as above and $t \in \mathbb{N}_0^2$ with $s_2 > t_1 \geq t_2$. Then*

$$\|\mathcal{J}(\mathcal{L}_K - \mathcal{L})\|_{H^s \to H^t} \leq C(1 + K^2)^{-\alpha/2}$$

for some $C > 0$ and $\alpha = s_2 - t_1$.

Proof. Let $\mathcal{L}'$ denote the transfer operator when considered on the larger space $H^t(\Omega)$. Then using the property $\mathcal{J}\mathcal{L} = \mathcal{L}'\mathcal{J}$, we have

$$\mathcal{J}(\mathcal{L}_k - \mathcal{L}) = \mathcal{J}P_K \mathcal{L} P_K - \mathcal{J}\mathcal{L} = (\mathcal{J}P_K - \mathcal{J})\mathcal{L}P_K - \mathcal{L}'(\mathcal{J}P_K - \mathcal{J}).$$

Thus,

$$\|\mathcal{J}(\mathcal{L}_K - \mathcal{L})\|_{H^s \to H^t} \leq (\|\mathcal{L}\|_{H^s \to H^s} \|P_K\|_{H^s \to H^s} + \|\mathcal{L}'\|_{H^t \to H^t}) \|\mathcal{J} - \mathcal{J}_K\|_{H^s \to H^t}$$
$$\leq C(1 + K^2)^{-\alpha/2},$$

where we have used Lemma 2.9, Lemma 3.1 and $\|P_K\|_{H^s \to H^s} \leq 1$. $\qquad \square$

Proposition 3.3. *Let $\mathcal{L}$ and the family $(\mathcal{L}_K)_{K \in \mathbb{N}}$ be as above, considered as operators on $H^s(\Omega)$ where $s \in \mathbb{N}^2$ with $s_1 \geq s_2$. Then, for $t \in \mathbb{N}_0^2$ with $s_2 > t_1 \geq t_2$ and for all $K \in \mathbb{N}$ we have*

$$\|(I - \mathcal{L}_K)^{-1} - (I - \mathcal{L})^{-1}\|_{H^s \to H^t} \leq C(1 + K^2)^{-\alpha/2},$$

for some $C > 0$ and $\alpha = s_2 - t_1$.

Proof. As $r(\mathcal{L}) \leq \|w\|_\infty < 1$ by Lemma 2.9, the operator $(I - \mathcal{L})^{-1}$ exists and is bounded. Let $(\mathcal{L}'_K)_{K \in \mathbb{N}}$ denote the family of transfer operators when considered on the larger space $H^t(\Omega)$. Similarly, as $\rho(\mathcal{L}'_K) \leq \|w\|_\infty < 1$ and the norms of $(\mathcal{L}'_K)^n$ are bounded uniformly in K by Lemma 2.9, the sums $\sum_{n=0}^\infty \|\mathcal{L}_K\|_{H^t \to H^t}^n$ are bounded by a constant independent of K and therefore $\|(I - \mathcal{L}_K)^{-1}\|_{H^t \to H^t}$ is uniformly bounded in K.

Using the property $\mathcal{J}(I - \mathcal{L}_K) = (I - \mathcal{L}'_K)\mathcal{J}$ and the second resolvent identity (see Equation (20)) we have

$$\|\mathcal{J}((I - \mathcal{L}_K)^{-1} - (I - \mathcal{L})^{-1})\|_{H^s \to H^t}$$
$$= \|(I - \mathcal{L}'_K)^{-1} \mathcal{J}(\mathcal{L}_K - \mathcal{L})(I - \mathcal{L})^{-1}\|_{H^t}$$
$$\leq \|(I - \mathcal{L}'_K)^{-1}\|_{H^t \to H^t} \|\mathcal{J}(\mathcal{L}_K - \mathcal{L})\|_{H^s \to H^t} \|(I - \mathcal{L})^{-1}\|_{H^s \to H^s}.$$

Using Proposition 3.2 for the bound on $\|\mathcal{J}(\mathcal{L}_K - \mathcal{L})\|_{H^s \to H^t}$ finishes the proof. $\qquad \square$

We are finally able to state and prove our main convergence result.

Theorem 3.4. *Let $\mathcal{L}$ and the family $(\mathcal{L}_K)_{K \in \mathbb{N}}$ be as above, considered as operators on $H^s(\Omega)$ with $s_1, s_2 \in \mathbb{N}$ and $s_1 \geq s_2 > t_1 \geq t_2 \geq 0$. Then for $f_0 \in H^s(\Omega)$ the operator equations (1) and (7) have unique solutions $f \in H^s(\Omega)$ and $f_K \in H^s(\Omega)$, respectively. Moreover there exist a constant $C > 0$ such that for all $K \in \mathbb{N}$ we have*

$$\|f_K - f\|_{H^t} \leq C(1 + K^2)^{-\alpha/2}\|f_0\|_{H^s},$$

where $\alpha = s_2 - t_1$.

Proof. The statement follows by writing $f = (I - \mathcal{L})^{-1}f_0$, $f_K = (I - \mathcal{L}_K)^{-1}f_0$ and using Proposition 3.3. $\qquad\square$

Remark 3.5. Note that for $f_0 \in \Lambda_K = P_K(H^s(\Omega))$, the unique solution f_K to (7) also lies in the finite-dimensional space Λ_K, so that (7) can be solved as a truly finite-dimensional problem.

4. Discussion and numerical experiments

Let us first summarise and rephrase our results in intuitive terms. Since the linear operator in Equation (1) fails to be compact, any finite-dimensional matrix representation would not reflect properties of the operator at all. Nevertheless the finite-dimensional representation in (7) provides a meaningful approximation for the solution of the inhomogeneous equation. For smooth periodic functions in location and angle of reflection, the solution of the approximated problem (7) converges to the solution of (1) in the Sobolev norm. The approximation error depends on the degree of smoothness of the inhomogeneous part. In addition, the approximation error is measured in a weaker norm, for instance the frequently used L^2 norm for the choice $t = (0, 0)$. The properties of this weaker norm also determine the speed of convergence. Broadly speaking, the convergence rate obeys a power law with the exponent being determined by the smoothness of the energy source and the norm used to measure the approximation error.

A finite amount of dissipation is a crucial ingredient in the entire approach, that is, the weight w has to satisfy $\|w\|_\infty < 1$. The simplest choice of a constant weight, $w(x, y) = \mu < 1$, corresponds to a dissipation which occurs at each collision at the boundary, for example, an attenuation of the sound wave caused by an inelastic reflection at the boundary of the cavity. Proper modelling of the damping parameters involved is a crucial aspect of the method and is necessary to describe realistic problems accurately [17]. For example, a linear attenuation in the medium would result in a path-length dependent weight $w(x, y) = \exp(-2\mu \cos(y))$. This choice, however, does not obey the stipulated bound as orbits with angles close to $y = \pm\pi/2$ have arbitrarily small path length, and hence small dissipation between subsequent collisions. We could overcome this particular problem by restricting the angle of reflection to non-tangential collisions, that is, $y \in (-(1 - \epsilon)\pi/2, (1 - \epsilon)\pi/2)$ for a small $\epsilon > 0$, effectively

constraining the permitted type of energy source. This however requires changing the Hilbert space and the projection operators, as the validity of $c_k(D_y^m f) = (\mathrm{i}2k)^m c_k(f)$ and $D_y P_k = P_k D_y$ is no longer given for a smooth function f on an interval instead of on a circle. One suitable choice could be the space of functions in $H^s(\Omega_\epsilon)$ with vanishing weak derivatives $D^\nu f$ on the boundary. A suitable basis is then the basis of Daubechies wavelets [30].

Of course, a circular billiard is very special, exhibiting a smooth boundary and a simple, analytically available collision map, making a rigorous treatment via Fourier analysis possible. The next simplest billiard geometry is that of an ellipse. While still an integrable smooth billiard with an explicit, albeit more complicated collision map, the existence of hyperbolic periodic orbits imposes restrictions on either the choice of function space or admissible weight functions (note that the spectral radius of the unweighted composition operator, considered, for example, on $H^{(1,1)}(\Omega)$, may depend on the maximal expansion rate in the system). Moreover, without recourse to a technical shortcut like Lemma 2.4, proving an analogue of Lemma 2.9 in this setting might be more involved. In the case of chaotic billiards, hyperbolic orbits are abundant, and hyperbolicity necessitates the use of more complicated anisotropic function spaces to account for expanding and contracting directions, see, for example [5, 6, 15]. Another important class are polygonal billiards, particularly relevant for numerical computations in DEA. These systems have zero Lyapunov exponents and the presence of discontinuities necessitates the use of function spaces including discontinuous functions, such as Sobolev spaces of low regularity or spaces of bounded variation [25]. An approximation scheme based on Fourier analysis is less suited to these systems, and Ulam-type discretisation schemes appear to have more potential.

To illustrate the impact of Theorem 3.4, we perform numerical simulations of circular billiards with constant damping $w(x, y) = \mu$. As a proxy for the error estimate we use the distance between approximations of subsequent order $\|f_{K+1} - f_K\|_{H^t}$, which obeys essentially the same upper bound

$$\|f_{K+1} - f_K\|_{H^t} \le \|f_{K+1} - f\|_{H^t} + \|f_K - f\|_{H^t} \le 2C(1 + K^2)^{-\alpha/2}\|f_0\|_{H^s}. \qquad (21)$$

Strictly speaking we have established this bound for integer vales of t_ℓ and s_ℓ only. With a little more effort this could be remedied by appealing to interpolation theory [29]. For simplicity of exposition we shall not pursue this here. For our numerical considerations we take the liberty to apply the bound above for non-integer values. For the norm $\|\cdot\|_{H^t}$, which estimates the truncation error, we use the choices $t = (0, 0)$, that is, the L^2 norm, and $t = (1, 1)$, a norm which is just outside the set of exponents guaranteeing pointwise convergence.

The transfer operator's action on Fourier modes is given in Equation (12). In order to use it for a numerical test, we have to use a representation for all Fourier modes, see Equation (A.2). We show results for three different choices of the initial boundary density f_0. They have in common that their support is given by the rectangle

$$R = \{(x, y) : x \in [\pi/6, \pi/6 + 4\pi/3], y \in [-0.8, 1.2]\}.$$

In order to define the boundary densities, we will use variables scaled on this rectangle according to $\tilde{x} = (x - \pi/6)/(4\pi/3)$ and $\tilde{y} = (y + 0.8)/2$ which take values between zero and one on R.

- Case G: a discontinuous function with $f_0(x, y) = 1$ for $(x, y) \in R$. This function is contained in $H^{(1/2-\epsilon, 1/2-\epsilon)}(\Omega)$ for any small $\epsilon > 0$. For simplicity of exposition we will use, however, the value $s_2 = 1/2$ in the discussion of the numerical results below.

- Case W_1: a continuous function given by $f_0(x, y) = \sqrt{\tilde{x}(1 - \tilde{x})}\sqrt{\tilde{y}(1 - \tilde{y})}$ for $(x, y) \in R$. This function lies in $H^{(1-\epsilon, 1-\epsilon)}(\Omega)$ for any small $\epsilon > 0$. As before, we use the choice $s_2 = 1$ in the discussion below.

- Case W_2: a smooth function given by $f_0(x, y) = \left(\sqrt{\tilde{x}(1 - \tilde{x})}\sqrt{\tilde{y}(1 - \tilde{y})}\right)^3$ for $(x, y) \in \mathrm{supp}(f_0)$. This function lies in $H^{(2-\epsilon, 2-\epsilon)}(\Omega)$ for any small $\epsilon > 0$ and we use the choice $s_2 = 2$ in our discussion.

The boundary densities above have their support in the rectangle R and exhibit different degrees of smoothness. In particular, the common support only covers part of the phase space and excludes the angles $y = \pm\pi/2$. If we had chosen a boundary density with support including these angles, that is, including tangential collisions, and which was otherwise smooth, then the convergence would still be slow. In this case the rate of convergence of the Fourier based approximation scheme would still be dominated by the discontinuities which occur when the density is extended in the angle variable y in a periodic fashion. Hence, choosing the variable y in Equation (2) to be cyclic is rather natural for the approximation scheme discussed here. This cyclicity condition cannot be dispensed with as the approximation scheme is based on Fourier expansions. Other approximation schemes, for example, a two-dimensional Ulam method could be investigated without imposing such a cyclicity condition, but this scheme does not achieve a speed-up as a result of increased regularity of the boundary density.

The data shown in Figure 1 confirm the upper bound in Theorem 3.4. For the L^2 norm, $t_1 = t_2 = 0$, we observe, in each case, convergence at a rate which is slightly faster than the theoretical prediction $\alpha = s_2 - t_1$. The power law decay of the truncation error shows up for large values of K and the onset of this scaling region shifts towards larger values if the initial boundary density becomes smooth. This should not come as a surprise, since the resolution of higher order derivatives requires higher order Fourier modes. For the parameter at the boundary of point-wise convergence $t = (1, 1)$, we see that the discontinuous boundary density fails to converge in line with our theoretical predictions. While Theorem 3.4 does not guarantee convergence in case W_1 either, the numerical data suggest an extremely slow convergence which is still consistent with the upper bound estimate $\alpha = s_2 - t_1 = 1 - 1 = 0$. Finally, for the smooth boundary density (case W_2) we observe a convergence rate slightly faster than the theoretical prediction.

From a dynamical perspective, circular billiards are trivial since the billiard map (2) is an integrable twist map. In order to get an idea of how dynamical properties impact on convergence properties we show numerical results for a deformed circle billiard which
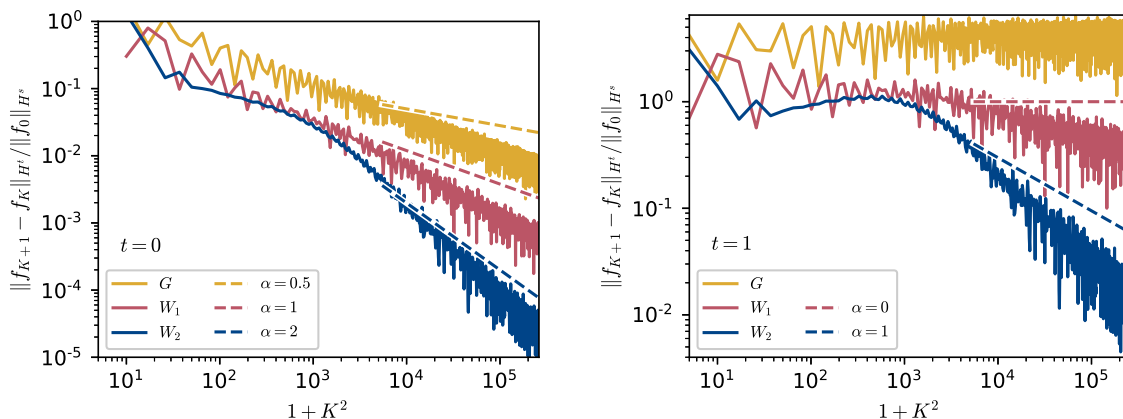
Figure 1. Error estimate $\|f_{K+1} - f_K\|_{H^t}$ for a circular billiard with constant damping $w(x, y) = \mu = 0.9$ as a function of the truncation order K on a double logarithmic scale. Left: $t_1 = t_2 = 0$ (convergence in L^2 norm), right: $t_1 = t_2 = 1$ (point-wise convergence, in essence). Results are displayed for three different initial boundary densities G: $s_2 = 1/2$ (yellow, top), W_1: $s_2 = 1$ (red, middle), W_2: $s_2 = 2$ (dark blue, bottom), see text. Lines show the power law decay according to Equation (21), $\alpha = s_2 - t_1$.

displays mixed regular and chaotic dynamics. For the deformation we choose the radius to depend on the polar angle x according to

$$r(x) = 1 + \delta \cos(mx), \tag{22}$$

where we choose $m = 3$ in the following. Deformations of this kind are known in the literature as Limaçon billiards [1]. We will cover the cases $\delta = 0.01$ and $\delta = 0.1$. For larger values, the billiard fails to be convex. In order to demonstrate the change in dynamical behaviour, Figure 2 shows the Poincare plot of the collision map T. For a small value of the deformation, $\delta = 0.01$, one still observes a fairly large number of invariant tori in accordance with general KAM folklore. The larger perturbation shown in Figure 2, $\delta = 0.1$, destroys most of the regular motion and renders the system chaotic with a few exceptions, for example, the highlighted period-3 island.

In order to calculate the convergence of the energy distribution we have to evaluate the matrix elements of the transfer operator. For the circular billiard, the only non-zero entries take the value $\pm\mu$ and follow the structure given by Equation (A.2). Once the circle has been deformed, the analytic calculation of the matrix elements is no longer possible. Even worse, the collision map is not given in closed analytic form either, so that an efficient numerical calculation becomes a nontrivial task (see the appendix for details). However, we are able to reduce the calculation of the matrix elements to double integrals with the kernel being given in closed analytic form, see Equation (A.3). Nevertheless, the numerical evaluation is still time consuming, in particular, since the matrix is no longer sparse. Hence, we can only calculate finite approximations up to $K = 30$. In order to reach the scaling regime (see Figure 1 for comparison) we employ
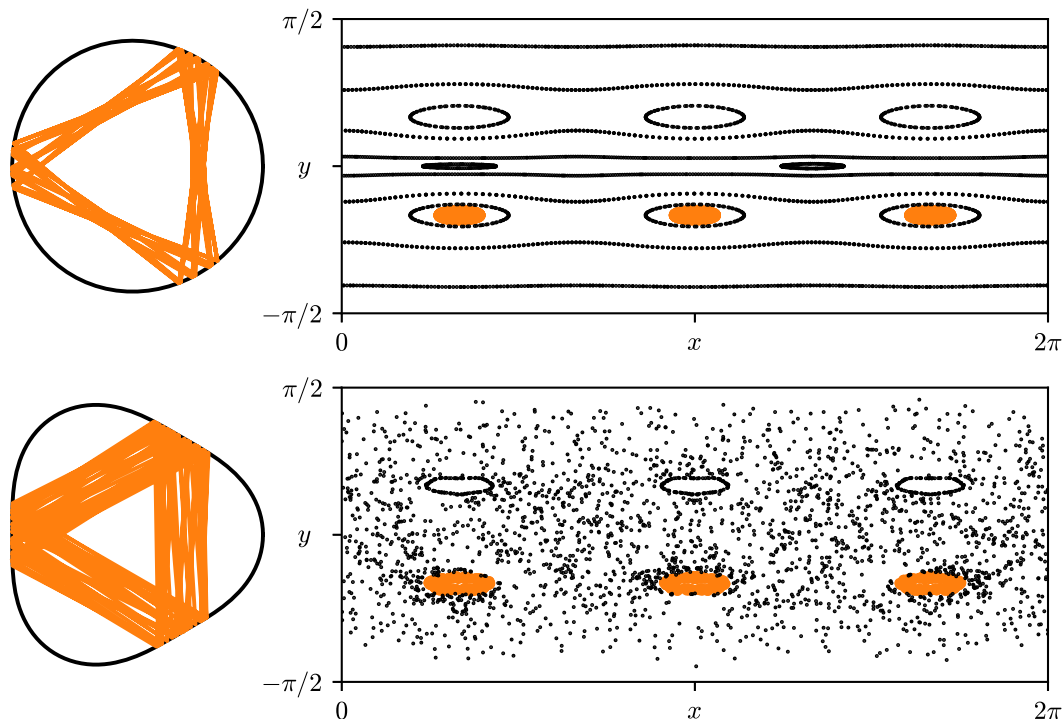
Figure 2. Billiard with orbit in configuration space (left) and Poincare plot of the boundary map T in the (x, y) phase space (right) for a deformed billiard according to (22). Top: weak deformation of the circle ($m = 3, \delta = 0.01$), bottom: strong but still convex deformation ($m = 3, \delta = 0.1$). The orbit depicted in real space is highlighted in phase space as well.

a stronger damping of $\mu = 0.1$. The results for the error measured in L^2 norm, that is, for the choice $t_1 = t_2 = 0$, are shown in Figure 3.

It is quite remarkable that the decay of the error is apparently almost unaffected by the degree of chaoticity. Hence the rigorous error estimate of Theorem 3.4 which covers the case $\delta = 0$ seems to have a wider range of applicability. While intuitively such an observation would not be surprising for nearly integrable cases it is quite counter-intuitive that the same error estimate may hold as well in strongly chaotic situations. However, our proof does not cover any of the deformed billiards and there does not seem to be an obvious way how the methodology can be generalised to these complicated cases. Nevertheless, it is reaffirming that our study of a simple dynamical system like the circular billiard has relevance for more complex dynamical behaviour.

Acknowledgments
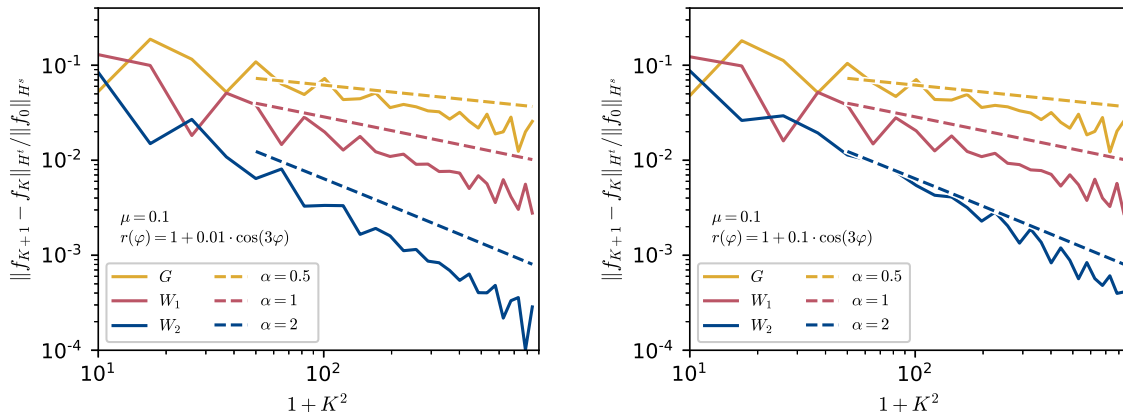
Figure 3. Error estimate $\|f_{K+1} - f_K\|_{H^t}$ in L^2 norm, $t = (0,0)$, for the energy density of a Limaçon billiard as a function of the truncation order K on a double logarithmic scale. Constant damping $w(x,y) = \mu = 0.1$ and two deformations, $\delta = 0.01$ (left) and $\delta = 0.1$ (right), are considered. Results are displayed for the three different initial boundary densities G: $s_2 = 1/2$ (yellow, top), W_1: $s_2 = 1$ (red, middle), W_2: $s_2 = 2$ (dark blue, bottom), see Figure 1. Lines indicate a power law decay, $\alpha = s_2 - t_1$, according to the rigorous estimate for circle billiards.

Appendix A. Matrix elements

Consider a convex billiard with boundary being given by $r(x)$ in polar coordinates where x denotes the polar angle (see, for example, Equation (22)). Denote by $(x', y') = (T_x(x,y), T_y(x,y))$ the collision map where x and x' label subsequent collisions with the boundary. Using a standard representation in terms of Fourier basis functions [28], the matrix elements $M_{l,k}$ of the transfer operator read

$$M_{l,k} = \frac{1}{2\pi^2} \int\limits_0^{2\pi} \int\limits_{-\pi/2}^{\pi/2} (\mathcal{C}_\phi e_k)(x,y) \,\overline{e_l(x,y)} \,\mathrm{d}y\mathrm{d}x$$

$$= \frac{1}{2\pi^2} \int\limits_0^{2\pi} \int\limits_{-\pi/2}^{\pi/2} \mathrm{e}^{\mathrm{i}k_1\phi_x(x,y) - \mathrm{i}l_1 x + 2\mathrm{i}k_2\phi_y(x,y) - 2\mathrm{i}l_2 y} \,\mathrm{d}y\mathrm{d}x$$

with $k = (k_1, k_2)$ and $l = (l_1, l_2)$.

In case of the perfect circle we get a representation which is given by a sparse matrix with only a few non-zero elements, close to the main diagonal, namely

$$(\mathcal{C}_\phi e_l)(x,y) = \sum_{k \in \mathbb{Z}^2} M_{l,k} \cdot e_k(x,y) \tag{A.1}$$

with the matrix elements

$$M_{l,k} = (-1)^{k_1} \,\delta_{k_1,l_1}\,\delta_{l_2,k_1+k_2} \qquad k,l \in \mathbb{Z}^2. \tag{A.2}$$

This is the extension of Equation (12) to all Fourier modes and it was used to calculate the values for Figure 1.

In order to eliminate the implicitly defined collision map we change integration variables from (x, y) to (x, x'). Using $y_1(x, x') = y$ and $y_2(x, x') = y'$ for the two scattering angles the matrix elements become

$$M_{l,k} = \frac{1}{2\pi^2} \int\limits_0^{2\pi}\int\limits_0^{2\pi} \left| \frac{\partial y_2(x, x')}{\partial x} \right| e^{i(k_1 x - l_1 x')} e^{2i(k_2 y_1(x,x') - l_2 y_2(x,x'))} dx' dx, \qquad (A.3)$$

where the additional factor is the Jacobian of the coordinate transformation. In contrast to the collision map T, the expressions $y_1(x, x')$ and $y_2(x, x')$ can be obtained in closed analytic form so that Equation (A.3) is easier to implement numerically.
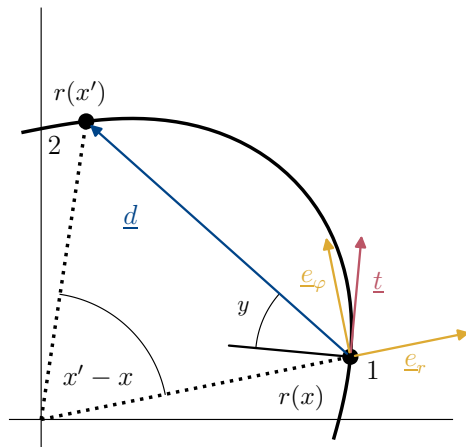
Figure A1. Geometric configuration of two subsequent collisions in a convex billiard with a particle moving from point 1 (with parameter value x) to point 2 (with parameter value x'). We also depict the ray vector $\underline{d}$, the tangent vector $\underline{t}$, and the unit vectors $\underline{e}_r$ and $\underline{e}_\varphi$ in polar coordinates.

Figure A1 shows a sketch of two subsequent collisions. The first scattering angle y_1 is given in terms of an inner product

$$\sin(y_1) = \underline{d} \cdot \underline{t} / (|\underline{d}||\underline{t}|) \,.$$

Since the position vector of the initial point is given by $r(x)\underline{e}_r$ the tangent is easily obtained as $\underline{t} = r'(x)\underline{e}_r + r(x)\underline{e}_\varphi$. The vector separating the two points of collision is given in terms of the local basis vectors by

$$\underline{d} = (r(x')\cos(x' - x) - r(x))\underline{e}_r + r(x')\sin(x' - x)\underline{e}_\varphi \,.$$

Hence the closed form expression for the first scattering angle reads

$$\sin(y_1) = \frac{r'(x)(r(x')\cos(x'-x) - r(x)) + r(x)r(x')\sin(x'-x)}{\sqrt{r^2(x) + (r'(x))^2}\sqrt{r^2(x) + r^2(x') - 2r(x)r(x')\cos(x'-x)}} \, . \tag{A.4}$$

The second scattering angle is obtained by interchanging the two points in Figure A1, that is, by swapping x and x' in Equation (A.4), and including an additional minus sign for the outgoing angle

$$\sin(y_2) = -\frac{r'(x')(r(x)\cos(x-x') - r(x')) + r(x')r(x)\sin(x-x')}{\sqrt{r^2(x') + (r'(x'))^2}\sqrt{r^2(x') + r^2(x) - 2r(x')r(x)\cos(x-x')}} \, .$$

References

[1] Bäcker A and Dullin H R, Symbolic dynamics and periodic orbits for the cardioid billiard, *J. Phys. A.* **30** (1997) 1991–2020

[2] Baladi V and Holschneider M, Approximation of nonessential spectrum of transfer operators, *Nonlinearity* **12** (1999) 525–538

[3] Baladi V, *Positive Transfer Operators and Decay of Correlations*, World Scientific, Singapore, 2000

[4] Baladi V, *Dynamical Zeta Functions and Dynamical Determinants for Hyperbolic Maps*, Springer, Cham, 2018

[5] Baladi V, Demers M F and Liverani C, Exponential decay of correlations for finite horizon Sinai billiard flows, *Inventiones mathematicae* **211** (2018) 39–177

[6] Blank M, Keller G and Liverani C, Ruelle-Perron-Frobenius spectrum for Anosov maps, *Nonlinearity* **15** (2002) 1905–1973

[7] Boyarsky A and Gora P, *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension*, Birkhäuser, Basel, 1997

[8] Chae K S and Ih J G, Prediction of vibrational energy distribution in the thin plate at high-frequency bands by using the ray tracing method, *J. Sound Vib.* **240** (2001) 263–292

[9] Chappell D J, Löchel D, Søndergaard N and Tanner G, Dynamical energy analysis on mesh grids: A new tool for describing the vibro-acoustic response of complex mechanical structures, *Wave Motion* **51** (2014) 589–597

[10] Chappell D J, Tanner G, Löchel D and Søndergaard N, Discrete flow mapping: transport of phase space densities on triangulated surfaces, *P. Roy. Soc. A* **469** (2013) 20130153

[11] Chen J and Wang H, Preasymptotics and asymptotics of approximation numbers of anisotropic Sobolev embeddings, *J. Complexity* **39** (2017) 94–110

[12] Dellnitz M and Junge O, On the approximation of complicated dynamical behavior, *SIAM J. Numer. Anal.* **36** (1999) 491–515

[13] Dellnitz M, Froyland G and Sertl S, On the isolated spectrum of Perron-Frobenius operator, *Nonlinearity* **13** (2000) 1171–1188

[14] Deschamps G A, Ray techniques in electromagnetics, *Proc. IEEE* **60** (1972) 1022–1035

[15] Gouëzel S and Liverani C, Banach spaces adapted to Anosov systems, *Ergod. Th. & Dyn. Sys.* **26** (2006) 189–217

[16] Haake F, *Quantum Signatures of Chaos*, 3rd Ed., Springer, Berlin Heidelberg, 2010

[17] Hartmann T, Morita S, Tanner G and Chappell D J, High-frequency structure- and air-borne sound transmission for a tractor model using Dynamical Energy Analysis, *Wave Motion* **87** (2019) 132–150

[18] Hartmann T, Tanner G, Xie G, Chappell D J and Bajars J, Modelling of high-frequency structure-borne sound transmission on FEM grids using the Discrete Flow Mapping technique, *J. Phys. Conf. Ser.* **744** (2016) 01237

[19] Keller G and Liverani, C, Stability of the spectrum for transfer operators, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **28** (1999) 141–152

[20] Klus S, Koltai P and Schütte C, On the numerical approximation of the Perron-Frobenius and Koopman operator, *J. Comp. Dyn.* **3** (2016) 57–79

[21] Kühn T, Sickel W and Ullrich T, Approximation numbers of Sobolev embeddings—Sharp constants and tractability, *J. Complexity* **30(2)** (2014) 95–116

[22] Liverani C, Rigorous numerical investigation of the statistical properties of piecewise expanding maps — A feasibility study, *Nonlinearity* **14** (2001) 463–490

[23] Mayer D H, *The Ruelle-Araki Transfer Operator in Classical Statistical Mechanics*, Springer-Verlag, Berlin, 1978

[24] Ruelle D, *Thermodynamic Formalism: the Mathematical Structures of Classical Equilibrium Statistical Mechanics*, Addison-Wesley, Reading, 1978

[25] Saussol B, Absolutely continuous invariant measures for multidimensional expanding maps, *Israel Journal of Mathematics*, **116** (2000) 223–248

[26] Savioja L and Svensson U P, Overview of geometrical room acoustic modeling techniques, *J. Acou. Soc. Am.* **138** (2015) 708–730

[27] Tanner G and Søndergaard N, Wave chaos in acoustics and elasticity, *J. Phys. A* **40** (2007) R443-R509

[28] Tanner G, Dynamical energy analysis—Determining wave energy distributions in vibro-acoustical structures in the high-frequency regime, *J. Sound. Vib.* **320** (2009) 1023–1038

[29] Triebel H, *Theory of Function Spaces III*, Birkhäuser, Basel, 1983

[30] Walker J S, *A Primer on Wavelets and their Scientific Applications*, CRC Press, Boca Raton, 2008