

Stable motions of high energy particles interacting via a repelling potential

V. Rom-Kedar* and D. Turaev†

September 1, 2022

Abstract

The motion of N particles interacting by a smooth repelling potential and confined to a compact d -dimensional region is proved to be, under mild conditions, non-ergodic for all sufficiently large energies. Specifically, choreographic solutions, for which all particles follow approximately the same path close to an elliptic periodic orbit of the single-particle system, are proved to be KAM stable in the high energy limit. Finally, it is proved that the motion of N repelling particles in a rectangular box is non-ergodic at high energies for a generic choice of interacting potential: there exists a KAM-stable periodic motion by which the particles move fast only in one direction, each on its own path, yet in synchrony with all the other parallel moving particles. Thus, we prove that for smooth interaction potentials the Boltzmann ergodic hypothesis fails for a finite number of particles even in the high energy limit at which the smooth system appears to be very close to the Boltzmann hard-sphere gas.

Contents

1	Introduction	2
2	Setup and Main results	7
2.1	Local assumptions on the single-particle system	8
2.2	Conditions on the coupling potential	12
2.3	Choreographic solutions of smooth multi-particle systems	14

*The Estrin Family Chair of Computer Science and Applied Mathematics, Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science (email:vered.rom-kedar@weizmann.ac.il).

†Department of Mathematics, Imperial College (email:d.turaev@imperial.ac.uk)

2.4	Choreographic solutions for repelling coupling potentials	15
2.5	High-energy particles in a bounded domain	18
2.6	Particles in a rectangular box	24
3	Smooth N-particle systems.	30
4	Mutually repelling particles in a container	42
4.1	A single particle at high energy.	43
4.1.1	The boundary layer dynamics	43
4.1.2	Flow-box coordinates in the boundary layer	49
4.1.3	The periodic orbit L_δ^*	51
4.1.4	Conditions SP1-SP3 for the orbit L_δ^*	54
4.2	Multi-particle system	56
5	Gas in a rectangular box	66
5.1	Action-angle coordinates near vertical periodic orbits.	66
5.2	The non-simultaneous impacts case	68
5.3	Smoothness properties of the averaged potential	70
5.4	The simultaneous impacts case	78
5.4.1	Expansion near the orbit $\mathbf{L}^*(\theta, \xi)$ of the uncoupled system.	79
5.4.2	Regularity of averaging	80
5.4.3	Normal form of the truncated averaged system.	81
5.4.4	Normal form of the full system.	85
6	Discussion	90
A	Choreographic solutions with equidistant phases.	93

1 Introduction

Can a large number N of repelling particles moving rapidly in a d (≥ 2)-dimensional domain D , remain forever bounded away from each other? We prove that such stable motion that avoids collisions occurs with positive probability. Borrowing the terminology from Celestial Mechanics [31, 10, 9, 16], the solutions we construct are of a choreographic type, i.e., the particles move essentially synchronously along the same path (or, along a family of parallel paths) with nearly constant phase shifts between them. It follows that systems of repelling particles are not ergodic, and have, in fact, KAM-stable states. In other words, for this class

of N -particle systems, the ergodic hypothesis of Boltzmann fails for any finite N in the high energy limit.

Establishing ergodicity of the Liouville measure (the Lebesgue measure restricted to a constant level of the Hamiltonian in the phase space) is a long-standing problem for conservative many-particle systems. The question is related to principal issues of the foundation of statistical mechanics, see e.g. [4, 24]. Classical statistical mechanics is based on the assumption, sometimes called *Gibbs postulate*, that macroscopic quantities describing the state of a large system of microscopic particles are averages over the Liouville measure in the phase space (the so-called micro-canonical ensemble). This postulate is supported by an overwhelming experimental evidence; the question is whether it can be inferred from the Hamiltonian formulation of dynamics by logical reasoning. Are there general properties of the Hamiltonian dynamics which make a general, yet finite system choose the Liouville measure over all other invariant measures?

The ergodicity of the Liouville measure could be such a property¹. However, by Kolmogorov-Arnold-Moser theorem, the ergodicity is violated for an open set of smooth Hamiltonians - for example, it is violated for energy levels near any non-degenerate minimum or maximum of the Hamiltonian function. Therefore, one cannot simply postulate ergodicity - it has to be justified by certain additional properties of the class of systems under consideration. Below we summarize some of the relevant works on the N -particle problem: proving ergodicity within the Sinai program of studying the (billiard) dynamics of the gas of hard balls, or, on the contrary, proving non-ergodicity by studying the emergence of stability islands. As we mentioned, at low energy, near local minima of the potential (i.e., near “ground states”) one expects, by KAM theory, that the system will be generically non-ergodic. Therefore, the mathematical question is to study ergodic properties of many-particle systems at high energies.

Hard spheres in a container: The idea going back to Boltzmann is that one can neglect the interactions between particles when the potential energy of the interaction is much smaller than their kinetic energy. This means that in the gas of sufficiently energetic particles, the particles motion is essentially free except for the short instances when the distance between some particles becomes small enough to create a strong repulsion force resulting in the fast change of the momenta. In the limit, one obtains the *Boltzmann gas* of N -hard spheres of diameter ρ , which interact only via momentarily elastic collisions and are

¹One may argue that macroscopic quantities are, in fact, time-averages, so they are indeed equal to the averages over the Liouville measure for a full-measure set of initial conditions when the Liouville measure is ergodic, by Birkhoff-Khinchin theorem.

confined to a d -dimensional container of the characteristic size L such that $N\rho^d \ll L^d$. This provides a *universal model* for any system of N particles in such a container for large values of the kinetic energy per particle, irrespective of the precise form of the repelling interaction potential.

Thus, proving the ergodicity of the Boltzmann gas – the Boltzmann-Sinai ergodic conjecture – is a corner-stone problem in the foundations of statistical mechanics. The Sinai program [46, 47, 48] was inspired by ideas of Krylov [26] and culminated in a series of works [41, 42, 43, 8, 44]. By this program, the ergodicity of the Boltzmann gas is inferred from the characteristic “Krylov-Sinai” instability of the elastic collision of spheres (or any convex bodies) in \mathbb{R}^d for $d \geq 2$: a small change in the momentum of the particle increases exponentially with the number of collisions. One can view the N -particle hard-sphere gas in d dimensions as a billiard in an Nd -dimensional domain [49]. The pair-wise collisions of the spheres correspond to boundaries of the domain – Krylov-Sinai instability means that these boundaries are (semi)-dispersing, which, for hard spheres moving on a flat torus or in a rectangular box, implies the hyperbolicity of the dynamics [8, 44, 41] and leads to the ergodicity of the Liouville measure [42, 43].

The Sinai program has led to seminal works in dynamical systems theory – it was one of the main sources for the development of ergodic theory of smooth dynamical systems, the theory of billiards and of general dynamical systems with singularities [21, 5, 11, 23]. However, it has also revealed the inherent difficulties in relating the Boltzmann gas dynamics to the problem of the ergodicity of multi-particle systems.

A well-recognized difficulty is the strong dependence of the hard-sphere dynamics on the container shape, see [6, 27, 25]. When the container boundary has a convex piece, the Nd -dimensional billiard representing the hard-spheres gas acquires a non-dispersing (focusing) boundary component, which makes the establishment of the hyperbolicity problematic. Notably, even in the case of a concave container, the ergodicity of the Boltzmann gas has been established only in a quite special geometrical set-up, for spheres of a sufficiently large diameter ρ [7].

The grander problem is the singularity of the hard-sphere system: the interaction potential jumps from zero to infinity when the distance between particles becomes equal to ρ . The Boltzmann gas serves as a universal limit of *smooth* multi-particle systems. Since this limit is singular, the question of which of its dynamical and statistical properties survive a regularization must be addressed.

Smooth billiard-like Hamiltonians provide a natural regularization of billiard dynamics. Such a Hamiltonian H is the sum of the kinetic energy term (a positive-definite

quadratic function of momenta) and a steep potential $V(q; \delta)$ associated with a billiard domain $D \subset \mathbb{R}^d$. The potential is a smooth function of $q \in D$ and it also depends on a small parameter δ (the inverse steepness), so that when $\delta \rightarrow 0$, the potential vanishes in the interior of D while staying bounded from below on the billiard boundary ∂D .

For example, in the present paper, we consider N -particle systems with a smooth interaction potential W which tends to $+\infty$ when the distance between the particles approaches ρ ; the particles are confined to an open bounded d -dimensional region by a smooth potential V which gets infinite on the boundary of this region. When we restrict the system to the energy surface $H = \frac{1}{2}Nh$ and scale the momenta by \sqrt{h} (so the energy is scaled by h), we obtain a billiard-like system with the steep potential $\delta(V + W)$, where $\delta = 1/h$; see the precise setup in Section 2.4. The limit $\delta \rightarrow 0$ for the fixed value of the rescaled energy $H = \frac{1}{2}N$ corresponds to the high energy limit of the unscaled system.

In [37, 52, 40, 51], we described a large class of billiard-like Hamiltonians with steep potentials that satisfy some natural growth and smoothness conditions. We proved for this class that the limit billiard dynamics which are represented by *regular* orbits – i.e., those which hit the billiard boundary ∂D away of its singularities and at angles away from zero – persist for sufficiently small δ . Namely, near the regular orbits, the local return maps of the smooth Hamiltonian flows to cross-sections that are bounded away from ∂D tend *with all derivatives* to those of the billiard as $\delta \rightarrow 0$ [37, 40, 51]. This implies that regular uniformly-hyperbolic sets and KAM-nondegenerate elliptic orbits of the billiard persist for sufficiently small δ in the smooth billiard-like system [37, 40].

On the other hand, we also showed that the regularization of dispersing billiards changes drastically their dynamics near *singular* orbits, such as orbits which are tangent to ∂D or which enter corner points in ∂D . Namely, the inherent hyperbolic structure of dispersing billiards *cannot survive the regularization* [51, 40]. In particular, singular periodic orbits of dispersing billiards give rise to stable periodic motions – hence to non-ergodic behavior – in the smooth system at arbitrarily small δ . Indeed, we proved, under quite general conditions, the loss of ergodicity due to the regularization for two-dimensional dispersing billiards [52, 51] and also for billiards with specific types of corners in any dimension [36, 38]. Applying this logic to the billiard that represents the Boltzmann gas, one concludes that the same dispersing geometry that creates the Krylov-Sinai instability of the colliding spheres is also responsible for the destruction of the associated hyperbolic structure – when the hard-spheres model is replaced by a more realistic model of particles interacting via a smooth potential. It is thus natural to conjecture that orbits of the system of N hard spheres which undergo sufficiently many instances of brushing (zero angle) collisions between the spheres or end at multi-collision points (simultaneous collisions of more than 2 particles) can produce islands of

stability of the corresponding system of N smoothly interacting particles at sufficiently high energy (see, e.g., discussion in [38]). However, proving the non-ergodicity of the N -particle problem in this fashion was not realized yet.

In this paper, we explore a new and different mechanism of the ergodicity loss of the hard-spheres system due to the smoothing. We establish the existence and stability of choreographic solutions for which highly-energetic particles, placed on the same periodic path or parallel paths, never come close to collisions. One can find such motions in the hard-spheres system as well (just take the diameter ρ small enough to ensure the spheres on the same path do not overlap and let them move with the same speed). However, they are unstable, as small discrepancies in the speed eventually lead to collisions of the spheres. As we show, if the speed of the synchronous motion of the particles is sufficiently high, a generic smoothing of the repulsive interaction potential stabilizes such type of solutions for some discrete set of particles phases.

The N -body problem of Celestial Mechanics has much in common with the N -particle problem discussed here, with the difference that the N -body problem usually refers to attracting interactions. In both cases, the pairwise interaction decays at large distances, whereas, at very small distances, the interaction potential is singular (see e.g. [15]). The full characterization of the mixed phase space dynamics for $N \geq 3$ is intractable due to non-integrability [22, 30, 50, 18]. Thus, finding special type of solutions, in particular KAM-stable periodic motions, is an achievement for such problems [3, 12, 13, 34, 45]. Choreographic solutions for the N -body problem were found by fixing the phase shift between the bodies to be constant and utilizing symmetries to establish that such solutions minimize the action [54, 10, 9, 32, 33, 17]. The avoidance of collisions for the attracting potential case follows from the observation that the action becomes infinite for sufficiently strong singularities (in particular, the Newtonian potential is not included). Numerically, the choreographic solutions with small N were found by continuation schemes also for the Newtonian potential [9, 35] and for the Lennard-Jones potential [16]. In these works, the choreographic solution is not induced by an external field, it is a genuine outcome of the particles interaction. Here, we propose to study choreographic solutions that follow a path dictated by a common background potential that governs the uncoupled dynamics. The method we use for proving the existence and stability of choreographic solutions is based on averaging. While we mostly focus on repelling potentials, the case of attracting potentials is covered by our scheme too, see Remarks 1 and 2 in Section 2.

The paper is ordered as follows. In Section 2 we list our main results regarding the existence and stability of choreography-type solutions in a system of N interacting particles in

a common background potential. We consider four different settings. Theorem 1 states that such KAM-stable motion exists in the case of N identical, weakly interacting particles when the particles are subject to a smooth background potential which admits a non-degenerate elliptic periodic orbit. Theorem 2 states that under some additional conditions the same result applies when the interaction potential is repelling - i.e. it diverges to $+\infty$ when the particles get close to each other. Theorem 3 states that the same result applies in the high energy limit for interacting particles with a repelling interaction potential when the background potential is billiard-like and the limiting billiard admits a KAM non-degenerate periodic orbit. Finally, Theorems 4 and 5 state that under some explicit non-degeneracy conditions there exist KAM-stable periodic motions of N repelling particles in a rectangular box. Sections 3 -5 contain the proofs of these theorems. Following the discussion section, the appendix establishes, by applying the results of [39], the existence of KAM-tori in Fermi-Pasta-Ulam type chains, which naturally arise in the averaging of identical particles systems.

2 Setup and Main results

By a particle, we mean a d degrees of freedom, autonomous Hamiltonian system with a Hamiltonian function $H_0(q, p)$, i.e.,

$$\dot{q} = \partial_p H_0(q, p), \quad \dot{p} = -\partial_q H_0(q, p), \quad (q, p) \in \mathbb{R}^{2d}. \quad (1)$$

Let this system have, at a certain energy value $H_0 = E^*$, an elliptic periodic orbit L^* with period $T = \frac{2\pi}{\omega_0}$. Let the equation of L^* be $(q, p) = (q^*(\omega_0 t), p^*(\omega_0 t))$ where (q^*, p^*) are some 2π -periodic functions.

Consider the system of N identical particles, each controlled by the same Hamiltonian H_0 , and allow the particles to interact with each other. We define the system of interacting particles by the Hamiltonian

$$H = \sum_{n=1, \dots, N} H_0(q^{(n)}, p^{(n)}) + \delta \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W(q^{(n)} - q^{(m)}) \quad (2)$$

where $(q^{(n)}, p^{(n)})$ are the coordinates and momenta of the n -th particle, $\delta \geq 0$ is a small coupling parameter, and W is the interaction potential. We assume that the particles are identical, so the interaction potential is the same for any pair of particles; similar results hold true also when the pairwise interaction potentials vary from pair to pair. Note that in (2) we sum over each pair twice and take the interaction to depend only on the difference

between the particles coordinates. With this choice, the potential is translation invariant and even. Such properties are natural from a physical point of view, yet, mathematically, they make the question of genericity more delicate. The proofs of Theorems 1 - 4 do not use these properties (rather, overcome them) so these theorems remain valid for arbitrary pair-wise interaction potential.

For $\delta = 0$, Hamiltonian (2) describes the motion of N non-interacting particles. It has “choreography” type solutions, for which each particle moves along the same periodic path L^* with a given phase shift:

$$\mathbf{L}^*(\theta) = \{q^{(n)} = q^*(\omega_0 t + \theta^{(n)}), \quad p^{(n)} = p^*(\omega_0 t + \theta^{(n)}), \quad n = 1, \dots, N\} \quad (3)$$

for an arbitrary set of fixed phases $\theta = (\theta^{(1)}, \dots, \theta^{(N)}) \in \mathbb{T}^N$. Below, we formulate conditions which ensure that for sufficiently small $\delta > 0$ choreographic motions persist for all time: the particles, modulo small oscillations, perpetually orbit L^* with the same frequency and with certain individual phase shifts θ .

The “equilibrium phases” θ are found as minima of the interaction potential averaged over the synchronous collective motion of the particles along L^* . We first perform the averaging for the case of uniformly bounded, smooth (C^∞) potential W ; after that we generalize the results to the case of repelling potentials, i.e., those which tend to $+\infty$ as $q^{(n)} - q^{(m)} \rightarrow 0$. Then we consider high-energy particles in a container of a generic shape. This corresponds to the single-particle Hamiltonian H_0 depending on δ in a singular way – the limit motion is a billiard in the domain where the particles are confined. The singularity in H_0 requires amendments to the averaging procedure and, also, additional conditions on the interaction potential for the persistence of choreographic motions along an elliptic periodic orbit of the billiard. Finally, we consider the special case of interacting high-energy particles in a rectangular box. The limit billiard does not have elliptic orbits in this case, however we show that a generic repelling interaction stabilizes choreographic motions along parabolic periodic orbits in the box.

2.1 Local assumptions on the single-particle system

First, we impose non-degeneracy conditions on the periodic orbit L^* of the one-particle system (1). We call these conditions single-particle (SP) assumptions. Recall that elliptic orbits exist in families parameterized by energy $E = H_0(q, p)$. Thus, by the assumption that the one-particle system (1) has at energy E^* an elliptic periodic orbit L^* , it follows that it has a smooth family $(q, p) = (q(t, E), p(t, E))$ of elliptic periodic orbits $L(E)$ such that the energy value $E = E^*$ corresponds to the original periodic orbit $L^* = L(E^*)$.

SP1: Acceleration assumption. *The period of $L(E)$ decreases with energy.*

Note that this assumption holds for billiard-like potentials [37, 40, 51], geodesic flows, and other settings where higher energy corresponds to a higher speed of the motion along the same (or almost the same) path in the configuration space.

Let us introduce symplectic coordinates (I_0, θ, z) where $I_0 \in \mathbb{R}$, $\theta \in \mathbb{S}^1 = \mathbb{R}^1/2\pi\mathbb{Z}$, $z \in \mathbb{R}^{2(d-1)}$, such that the surface filled by the periodic orbits $L(E)$ is given by $z = 0$.² Moreover, we choose (I_0, θ) such that they give action-angle variables for system (1) restricted to the surface $z = 0$. This means that on any of the periodic orbits $L(E)$ that foliate this surface the value of I_0 stays constant and equal to the signed area between $L(E)$ and $L(E^*)$ (and the variable θ is symplectic conjugate to I_0). Thus, the Hamiltonian restricted to this surface is a function of I_0 only and the frequency of the periodic orbit $L(E^*)$ is equal to

$$\omega_0 = \partial_{I_0} H_0(I_0, \theta, 0)|_{H_0=E^*}.$$

Thus, Assumption SP1 reads as

$$a = \partial_{I_0 I_0} H_0(I_0, \theta, 0)|_{H_0=E^*} > 0. \quad (4)$$

Let the multipliers³ of L^* be $e^{\pm i \frac{2\pi}{\omega_0} \omega_1}, \dots, e^{\pm i \frac{2\pi}{\omega_0} \omega_{d-1}}$.

SP2: Non-resonance assumption. *The frequencies $\omega = (\omega_0, \dots, \omega_{d-1})$ are not in a strong resonance:*

$$m_0 \omega_0 + \sum_{j=1}^{d-1} m_j \omega_j \neq 0 \quad (5)$$

for every integer m_0 and every integer vector (m_1, \dots, m_{d-1}) such that $1 \leq \sum_{j=1}^{d-1} |m_j| \leq 4$.

By this assumption, system (1) can be brought to Birkhoff normal form up to fourth

²This is done as follows: one first straightens this surface by a smooth coordinate transformation which is not necessarily symplectic; this transformation may change the symplectic form – then one brings the symplectic form back to the standard form by a smooth transformation which leaves the surface $z = 0$ invariant.

³Recall that the multipliers of the periodic orbit L^* are defined as follows. Consider a restriction of the system to the $(2d - 1)$ -dimensional energy level $H_0 = E^*$, take a $(2d - 2)$ -dimensional cross-section to L^* in this energy level, and consider the Poincaré map (the map defined by the orbits of the system) on the cross-section. The intersection point of the periodic orbit with the cross-section is a fixed point of the map. The eigenvalues of the linearization matrix of the Poincaré map at this point are called the multipliers of L^* . Since L^* is elliptic, all its multipliers are not real and lie on the unit circle.

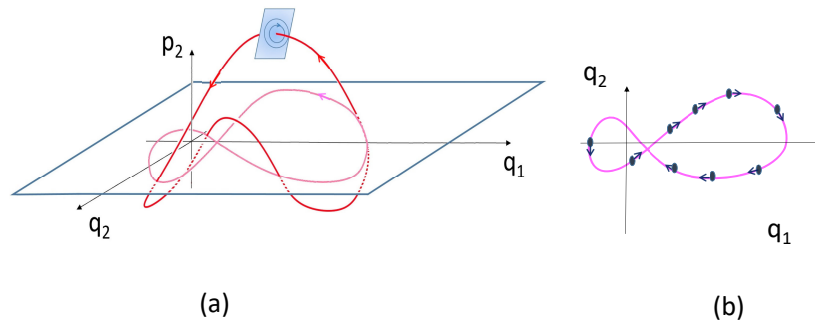


Figure 1: (a) An elliptic periodic orbit of a 2 degrees of freedom system on the fixed energy level $H_0 = E^*$. The elliptic periodic orbit L^* (red curve) is surrounded by KAM tori - the blue curves correspond to the intersection of these tori with a transverse cross-section. The pink curve shows the projection of L^* to the two-dimensional configuration space - here the projected curve has a single self-intersection point. (b) Theorem 1 shows that for a generic smooth interaction potential, any number of weakly interacting particles can orbit the same path in the configuration space. Theorem 2 shows that the multi-particle choreography is KAM-stable also for repelling potentials.

order [3]. Namely, in a sufficiently small neighborhood of L^* one can perform a symplectic coordinate transformation

$$(q, p) = (\hat{q}(I_0, \theta, z), \hat{p}(I_0, \theta, z))$$

where $(I_0, \theta, z) \in \mathbb{R} \times \mathbb{S}^1 \times \mathbb{R}^{2(d-1)}$, such that the Hamiltonian H_0 takes the form

$$H_0(I_0, \theta, z) = E^* + \omega I + \frac{1}{2} I^\top A I + g(I_0, \theta, z), \quad (6)$$

where $g(I_0, \theta, z) = g_0(I_0) + g_1(I_0)\hat{I} + O(\|z\|^4|I_0| + \|z\|^5)$ with $g_0 = O(I_0^3)$, $g_1 = O(I_0^2)$. Here we use the notation $z = (z_1, \dots, z_{d-1})$ where $z_j = (x_j, y_j) \in \mathbb{R}^2$, $I = (I_0, \hat{I})$, where $\hat{I} = (I_1, \dots, I_{d-1})$ and $I_j = \frac{1}{2}z_j^2$ denote the actions in the directions transverse to the orbit L^* , $j = 1, \dots, d-1$. We think of I as a column vector, the frequency vector, ω , is the row vector with components $\omega_0, \omega_1, \dots, \omega_{d-1}$, and A is a symmetric $d \times d$ matrix with constant coefficients. We denote

$$A = \begin{pmatrix} a & b \\ b^\top & \hat{A} \end{pmatrix}, \quad (7)$$

where a is a scalar (it is given by (4) and is strictly positive by Assumption SP1), $b \in \mathbb{R}^{d-1}$ is a row vector, and \hat{A} is a symmetric $(d-1) \times (d-1)$ matrix with elements $\{a_{kj}\}_{k,j=1,\dots,d-1}$.

The system of differential equations defined by the Hamiltonian (6) has the form

$$\begin{aligned} \dot{I}_0 &= O(\|I\|^{5/2}), & \dot{\theta} &= \omega_0 + aI_0 + \sum_{j=1}^{d-1} b_j I_j + O(I^2), \\ \dot{x}_j &= (\omega_j + b_j I_0 + \sum_{k=1}^{d-1} a_{kj} I_k) y_j + O(I^2), \\ \dot{y}_j &= -(\omega_j + b_j I_0 + \sum_{k=1}^{d-1} a_{kj} I_k) x_j + O(I^2), & j &= 1, \dots, d-1. \end{aligned}$$

To the main order, the motion is a nonlinear rotation – the rotation of the phase θ corresponds to the motion along the orbit L^* (the circle $I = 0$ in these coordinates), and the rotation of (x_j, y_j) describes the transverse oscillations. The frequencies of the oscillations depend on the actions I_0, \dots, I_{d-1} , and we assume that this dependence is non-degenerate:

SP3: Twist assumption. *The orbit L^* satisfies the twist condition and the iso-energetic twist condition:*

$$\det A = \det \begin{pmatrix} a & b \\ b^\top & \hat{A} \end{pmatrix} \neq 0 \quad (8)$$

and

$$\det A_\omega = \det \begin{pmatrix} 0 & \omega_0 & \hat{\omega} \\ \omega_0 & a & b \\ \hat{\omega}^\top & b^\top & \hat{A} \end{pmatrix} \neq 0, \quad (9)$$

where $\hat{\omega} = (\omega_1, \dots, \omega_{d-1})$.

This completes the list of assumptions on the single-particle motion, assuring that L^* is surrounded by KAM tori.

2.2 Conditions on the coupling potential

Next, we impose a non-degeneracy condition on the coupling potential in the multi-particle system (2). We call such conditions Interacting Particles (IP) assumptions. Consider the motion of N uncoupled particles over the same periodic orbit L^* , as given by (3), and introduce the *averaged interaction potential*

$$U(\theta^{(1)}, \dots, \theta^{(N)}) = \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W_{avg}(\theta^{(n)} - \theta^{(m)}), \quad (10)$$

where

$$W_{avg}(\theta^{(n)} - \theta^{(m)}) = \frac{1}{T} \int_0^T W(q^*(\omega_0 t + \theta^{(n)}) - q^*(\omega_0 t + \theta^{(m)})) dt. \quad (11)$$

Note that U is invariant under translations: $U(\theta^{(1)}, \dots, \theta^{(N)}) = U(\theta^{(1)} + c, \dots, \theta^{(N)} + c)$ for any constant c . As U is a continuous function on the torus \mathbb{T}^N , it must have a point of minimum. By the translation invariance, the minima of U form lines in \mathbb{T}^N :

$$\theta^{(n)} = \theta_{min}^{(n)} + c, \quad n = 1, \dots, N. \quad (12)$$

We take any such line; by varying c , we can always make $\theta_{min}^{(1)} + \dots + \theta_{min}^{(N)} = 0$ in (12). Introduce coordinates $(\varphi, \psi_1, \dots, \psi_{N-1})$ in a small neighborhood of this line on \mathbb{T}^N such that $\varphi = \frac{1}{N}(\theta^{(1)} + \dots + \theta^{(N)})$ and the transformation $(\theta^{(1)} - \theta_{min}^{(1)}, \dots, \theta^{(N)} - \theta_{min}^{(N)}) \mapsto (\varphi, \psi_1, \dots, \psi_{N-1})$ is linear and, up to the factor $\frac{1}{\sqrt{N}}$, orthogonal. In the new coordinates, the line of minima is the line $\psi = 0$ and the averaged potential U is independent of φ , so we denote

$$\hat{U}(\psi) = U(\theta^{(1)}(\varphi, \psi), \dots, \theta^{(N)}(\varphi, \psi)). \quad (13)$$

We show in Section 3 that for small δ , the evolution of phases of interacting particles moving along the path L^* is, to the main order, governed by the potential U . We are looking for stable motions, therefore, we make the following assumption on the interaction potential.

IP1: KAM assumption. *The local minimum of $\hat{U}(\psi)$ at the origin is non-degenerate;*

namely, the eigenvalues of the Hessian matrix $\left(\partial_{\psi_j\psi_k}\hat{U}\right)_{j,k=1,\dots,N-1}$ are strictly positive, so the origin is an elliptic equilibrium point for the system defined by the Hamiltonian

$$H = \frac{a}{2N}J^2 + \hat{U}(\psi), \quad J \in \mathbb{R}^{N-1}. \quad (14)$$

Moreover, in any neighborhood of the equilibrium this system has KAM-tori.

In view of the symmetries of $\hat{U}(\psi)$ we explain this assumption in more details. The eigenvalues of the Hessian matrix $\left(\partial_{\psi_j\psi_k}\hat{U}\right)_{j,k=1,\dots,N-1}$ are the squares of the frequencies of the small oscillations near the equilibrium configuration of the phases. The existence of KAM-tori (by definition they form a set of positive measure) is, for example, achieved when there are no resonances up to order four between the frequencies of small oscillations, and the corresponding Birkhoff normal form satisfies the twist condition. This is a natural requirement for a generic potential \hat{U} .

However, since we consider the system of identical particles, the averaged potential $U = \sum_{m \neq n} W_{avg}(\theta^{(n)} - \theta^{(m)})$ is symmetric with respect to any permutation of the phases θ (in particular, since the sum is made over all pairs of different phases, we can always think of W_{avg} as an even function). If the line (12) of minimum of U is not preserved by the symmetry, then, as in the general case, the absence of small resonances and twist suffice for the existence of KAM tori. On the other hand, when the line (12) is symmetric with respect to some permutation of phases, the potential \hat{U} inherits this symmetry, which may lead to resonances. In this case, the question of the existence of KAM-tori in system (14) cannot be reduced to the standard genericity assumptions and has to be specially addressed.

A natural example⁴ is given by the equidistant distribution of phases: $\theta^{(n)} = 2\pi \frac{n}{N}$. It is easy to see that the gradient of U vanishes for this choice of θ 's and, moreover, U indeed has a minimum at such configuration if, for example, the second derivative of W_{avg} is positive at the points $2\pi \frac{j}{N}$, $j = 1, \dots, N - 1$ (which is natural for repelling forces). The corresponding line of minima (12) is symmetric with respect to the cyclic permutation $(\theta^{(1)}, \dots, \theta^{(N)}) \mapsto (\theta^{(2)}, \dots, \theta^{(N)}, \theta^{(1)})$; this is responsible for the unavoidable creation of strong resonances between the frequencies of small oscillations in system (14), similarly to the Fermi-Pasta-Ulam (FPU) chain, see Appendix A. In this case one cannot bring the system to the Birkhoff normal form. Yet, as explained in Appendix A, one can generalize the theory that Rink built for the FPU [39] and show that system (14) has KAM-tori near the minimum of \hat{U} for a generic pairwise interaction potential W_{avg} (i.e., Assumption IP1

⁴A simpler example is the symmetric configuration $\theta^{(1)} = \dots = \theta^{(N)}$ (all the particles are at the same point), yet it is not very relevant in our setting – we are interested in the case of repelling potentials, so having all particles glued together for all time should not give a minimum of the averaged potential.

holds generically for the minimum line of U corresponding to the equidistant distribution of phases). The genericity assumption may be checked by calculating the coefficients of the Rink normal form. Thus, even though Assumption IP1 could be difficult to check in general (system (14) has $(N - 1)$ degrees of freedom, which can be arbitrarily large), it is not very restrictive.

2.3 Choreographic solutions of smooth multi-particle systems

The next theorem establishes the existence of KAM-stable choreographic solutions in the multi-particle system (2) near $\mathbf{L}^*(\theta_{min})$, the choreographic solution (3) of the uncoupled system, where θ_{min} is a minimum of the averaged potential. For a positive measure set of initial conditions, all particles follow approximately the same path with the phase difference between particles n and m remaining close to $\theta_{min}^{(n)} - \theta_{min}^{(m)}$ for all time.

Theorem 1. *Consider the system (2) where the single-particle system (1) has a periodic orbit L^* satisfying the acceleration (SP1), no-resonance (SP2), and twist (SP3) assumptions, and let the C^∞ -smooth bounded pairwise interaction potential $W(q)$ be such that its averaged interaction potential \hat{U} admits a minimum satisfying the KAM Assumption (IP1). Then, for all sufficiently small $\delta > 0$, system (2) admits a positive measure set of initial conditions corresponding to quasi-periodic solutions which satisfy, uniformly for all time t ,*

$$(q^{(n)}(t), p^{(n)}(t)) = (q^*(\bar{\omega}t + \theta_{min}^{(n)}), p^*(\bar{\omega}t + \theta_{min}^{(n)})) + O(\delta^{1/4}), \quad n = 1, \dots, N, \quad (15)$$

with some constant $\bar{\omega}$ (which may depend on initial conditions) such that $\bar{\omega} = \omega_0 + O(\delta^{1/2})$.

The theorem is proven in Section 3.

Remark 1. The conclusion of Theorem 1 about the existence of a positive measure set of quasiperiodic choreographic motions also holds near non-degenerate *maxima* of \hat{U} , provided the acceleration assumption SP1 is reversed to deceleration, i.e., if the period of L^* *increases* with energy (as is the case near homoclinic loops, see Figure 3). Indeed, the quasiperiodic choreographic solutions remain such if we reverse the direction of time. This corresponds to changing the sign of the Hamiltonian (2), i.e., of both H_0 and W . The change of the sign of H_0 makes the period of L^* grow with energy (so the coefficient a in (14) becomes negative, cf. (4)), while changing the sign of W makes a minimum of the averaged potential a maximum.

2.4 Choreographic solutions for repelling coupling potentials

As the single-particle system (6) near the elliptic orbit is nearly-integrable, the multi-particle system (2) near the set of choreographic solutions (3) is also nearly-integrable for small δ . Therefore, the existence of KAM-tori established by the above theorem is not surprising. However, this result admits a generalization to systems with singularities, where the concept of near-integrability is not automatically applicable.

The simplest case corresponds to a singularity in the interaction potential.

Definition 2.1. We call a potential $W(q)$ **repelling**, if it is smooth and bounded from below for all $\|q\| > \rho \geq 0$ and $W(q) \rightarrow +\infty$ as $\|q\| \rightarrow \rho$. For $\|q\| \leq \rho$, the function W is infinite.

Note that the repulsion assumption is made only for small distances, at large distances the potential may be attracting (e.g. the theory applies to the Lennard-Jones potential). If the pairwise interaction potential W in (2) satisfies this definition, the perturbation term $\delta \sum_{n \neq m} W(q^{(n)} - q^{(m)})$ can become large for arbitrary small δ - this occurs if $q^{(n)}$ gets sufficiently close to $q^{(m)}$ for some $n \neq m$. In particular, the averaged potential W_{avg} in (11) may have singularities when two particles moving on the same path L^* collide. When $\rho = 0$, this happens when two phases are identical ($\theta^{(n)} = \theta^{(m)}$ for some $m \neq n$) or when the orbit L^* has self-crossings in the q -space, see 1 (i.e., $q^*(\theta^{(n)}) = q^*(\theta^{(m)})$ for some $\theta^{(n)} \neq \theta^{(m)}$). If the dimension d of the q -space is larger than 2, then a typical periodic orbit has no self-crossings. However, the existence of self-crossings is a robust phenomenon when $d = 2$, or, in higher dimensions, when there are certain symmetries.

In general, we define the *collision set* $\Theta_c \subseteq \mathbb{T}^N$ as the set of all initial phases $(\theta^{(1)}, \dots, \theta^{(n)})$ for which the motion of uncoupled particles along the same path L^* (see (3)) leads to a collision at a certain time t_c :

$$\Theta_c = \{(\theta^{(1)}, \dots, \theta^{(n)}) \mid \|q^*(\omega_0 t_c + \theta^{(n)}) - q^*(\omega_0 t_c + \theta^{(m)})\| \leq \rho \text{ for some } n \neq m \text{ and some } t_c \in [0, T]\}.$$

The collision set is closed. For $\rho = 0$, it is, typically, a union of codimension-1 hypersurfaces in \mathbb{T}^N , whereas for $\rho > 0$ this set has a non-empty interior, see Figure 2. For too large ρ it coincides with the torus. For $(\theta^{(1)}, \dots, \theta^{(n)})$ outside the collision set, the averaged potential $U(\theta^{(1)}, \dots, \theta^{(N)})$ is well-defined and is a smooth function, bounded from below. Therefore, if Θ_c is not the whole torus \mathbb{T}^N , then U attains a finite minimum.

IP2: No-collision assumption. *There exists a minimum line (of the form (12)) of the averaged potential $U(\theta^{(1)}, \dots, \theta^{(N)})$ which is collision free, i.e., this line does not intersect the collision set Θ_c .*

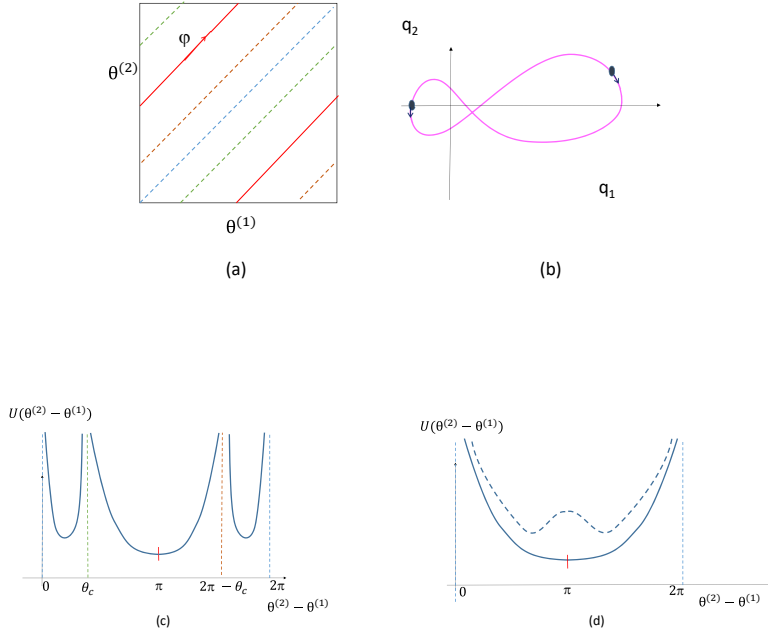


Figure 2: A two-particle system for an orbit with one self-intersection (a-c) and averaged potential U for the case of no self-intersection (d). (a) The collision set for $\rho = 0$: dashed blue line corresponds to the two particles coalescing, $\theta^{(1)} = \theta^{(2)}$; the dashed green and brown lines correspond to the two possible collisions at the intersection point, $\theta^{(1)} = \theta^{(2)} + \theta_c$ and $\theta^{(1)} = \theta^{(2)} - \theta_c$, respectively. (b) The two particles on the self-intersecting curve. (c) The averaged potential is infinite at the collision set. Provided the collision set does not include $\theta_c = \pi$, it must have at least 3 minima (it may attain additional minima points). As the averaged potential is even, π must be an extremal point, here a minimum. (d) If there is no self-crossing, the only singularity of U is at the blue line, so, if U has a single minimum, it must be at $\theta = \pi$. It is also possible that $\theta = \pi$ is a maximum of U (dashed curve), and then non-symmetric minima must exist.

If $\rho = 0$ and the elliptic periodic orbit $(q^*(\omega_0 t), p^*(\omega_0 t))$ of (1) has no self-intersection points, then Assumption IP2 automatically holds for any repelling potential and for any minimum line of U . Indeed, in this case the set Θ_c is exactly the set where at least two phases are equal. Then, at least two particles on the path L^* have the same coordinate q for all time, hence the corresponding integrand in (10) is infinite on the whole interval $[0, T]$, making the average potential U infinite everywhere on Θ_c . This cannot be a minimum of U .

In order to conclude the same when L^* has self-crossings or when $\rho > 0$, it is enough to have the repelling interaction sufficiently strong. For example:

Lemma 2.2. *If the pairwise interaction potential W satisfies, for $\|q\| > \rho$, the growth condition*

$$W(q) \geq \frac{C_1}{\|q\| - \rho} - C_2, \quad C_1 > 0, \quad (16)$$

with some constants $C_{1,2}$, then the no-collision assumption IP2 holds, unless the collision set coincides with the whole torus \mathbb{T}^N .

Proof. It is enough to show that the averaged potential U is infinite for all $(\theta^{(1)}, \dots, \theta^{(n)})$ from the collision set Θ_c . If $(\theta^{(1)}, \dots, \theta^{(n)}) \in \Theta_c$, then for some $n \neq m$ either $\|q^*(\omega_0 t + \theta^{(n)}) - q^*(\omega_0 t + \theta^{(m)})\| \leq \rho$ for all t , or there exists a value of $t = t_c$ such that $\|q^*(\omega_0 t_c + \theta^{(n)}) - q^*(\omega_0 t_c + \theta^{(m)})\| = \rho$ and $\|q^*(\omega_0 t + \theta^{(n)}) - q^*(\omega_0 t + \theta^{(m)})\| > \rho$ when t approaches t_c . In the first case, an integrand in (10) is infinite on the whole interval $[0, T]$, thus making U infinite. In the second case, since the derivative $\frac{d}{dt}q^*(\omega_0 t + \theta)$ is bounded, the distance between the particles decays at least linearly in $(t - t_c)$, so the integrand $W(q^*(\omega_0 t + \theta^{(n)}) - q^*(\omega_0 t + \theta^{(m)}))$ grows, by (16), proportionally to $(t - t_c)^{-1}$ or faster, hence the integral (11) diverges, i.e., U is infinite in this case as well. \square

There can, however, be cases when $\Theta_c = \mathbb{T}^N$ and Assumption IP2 does not hold. For example, if the system is reversible, then there can exist orbits (like period-2 orbits in billiards) for which $q^*(\omega_0 t) = q^*(2\pi - \omega_0 t)$ for all t . In this case, collisions are unavoidable even for $\rho = 0$. Note, however, that generically, since the orbit L^* is elliptic, around it one can find resonant periodic orbits with at most finitely many self-crossings, and the previous remarks apply for sufficiently small ρ .

The following result generalizes Theorem 1 to the case of repelling interaction potentials.

Theorem 2. *Consider system (2) where the single-particle system (1) has a periodic orbit L^* satisfying Assumptions SP1, SP2, SP3, and the C^∞ -smooth, repelling potential $W(q)$ is such that its averaged interaction potential \hat{U} admits a minimum satisfying the KAM assumption IP1 and the no-collision assumption IP2. Then, for all sufficiently small δ , the system*

admits a positive measure set of initial conditions corresponding to quasi-periodic solutions as in Theorem 1.

Proof. By Assumption IP2, we have that the uncoupled particles moving by the path L^* (see (3)) stay away from collisions, i.e., the distances $\|q^*(\omega_0 t + \theta_{min}^{(n)}) - q^*(\omega_0 t + \theta_{min}^{(m)})\|$ stay bounded away from ρ for any $m \neq n$ and for all t . Therefore,

$$W(q^*(\omega_0 t + \theta_{min}^{(n)}) - q^*(\omega_0 t + \theta_{min}^{(m)})) < K \quad (17)$$

for some constant K .

Replace the potential W by smooth and everywhere bounded potential W^{cut} which coincides with W when $W < K + 1$. The corresponding averaged potential U^{cut} coincides with U in a neighborhood of $(\theta^{(1)}, \dots, \theta^{(N)}) = (\theta_{min}^{(1)}, \dots, \theta_{min}^{(N)})$, i.e., U^{cut} has the same minimum $(\theta_{min}^{(1)}, \dots, \theta_{min}^{(N)})$. By Theorem 1 the multi-particle system (2) with potential W^{cut} has a positive measure set of quasiperiodic solutions for which $q(t)$ remains $O(\delta^{1/4})$ -close to $q^*(\bar{\omega}t + \theta_{min}^{(1)}), \dots, q^*(\bar{\omega}t + \theta_{min}^{(N)})$ where $\bar{\omega} = \omega_0 + O(\delta^{1/4})$. For sufficiently small δ , these solutions correspond to particles which stay away from each other for all times (because, for all times, $q^*(\omega_0 t + \theta_{min}^{(n)})$ are bounded away from each other, thus, the same is true for $q^*(\bar{\omega}t + \theta_{min}^{(n)})$). Hence, for such solutions, if δ is small enough, the potential W is bounded by (17), so W^{cut} coincides with W . Thus, they are also solutions of system (2) with the original potential W . \square

Remark 2. The result stays true if we reverse time, as in Remark 1 of Theorem 1. This means that we can replace the repelling potential W by a potential which *attracting*, i.e., bounded from above and tending to $-\infty$ as particles come close to each other. Then, Theorem 2 implies that a positive measure set of quasiperiodic choreographic motions exists with phases close to the *maximum* of the averaged potential \hat{U} (if it satisfies IP1 and IP2), provided the elliptic orbit L^* of the single-particle system satisfies the non-degeneracy assumptions SP2 and SP3, and the acceleration assumption SP1 is reversed to deceleration, see Figure 3.

2.5 High-energy particles in a bounded domain

Next, we apply the above methodology to the case of a system of repelling particles which are confined in a bounded domain. Let $D \subset \mathbb{R}^d$ be a domain with a smooth (C^∞) or piecewise smooth boundary ∂D (in the piecewise smooth case, we call the points where ∂D is smooth

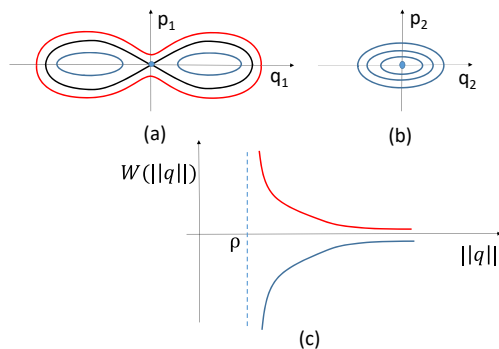


Figure 3: Accelerating and decelerating families of elliptic orbits coexist near homoclinic loops to a saddle-center. Consider a product system of a Duffing oscillator (a) and a center (b). Then, in the vicinity of the homoclinic loop to the saddle-center there exist three families of periodic orbits corresponding to $(q_2, p_2) = 0$. For the family of periodic orbits which are outside the separatrix (red curve in the (q_1, p_1) space) the period decreases with energy, whereas for the two families of periodic orbits which are inside the separatrix (blue curves in the (q_1, p_1) space) the period decreases with energy. By Theorem 1 and Remark 1, for a smooth interaction potential, choreographic solutions exist near the averaged potential minima lines for any fixed outer orbit, and near the potential maxima lines for any given inner orbit. For smooth potentials, both types of lines must exist. Similar conclusions apply to singular potentials, where here, when the potential is repelling (red curve in (c)) a minimum line must exist, whereas, when the potential is attracting (blue curve in (c)) a maximum line must exist, see Remark 2.

non-singular). A particle confined in D is described by the Hamiltonian:

$$H = \frac{p^2}{2} + V(q),$$

where the potential $V(q)$ is a C^∞ -function defined in the interior of D and tending to $+\infty$ on ∂D (here, if the particle is a ball of a finite diameter ρ , the domain D is the set of all possible positions of the ball center). When the growth of V at the approach to ∂D is reasonably regular, the high-energy motion limits to the billiard in D , as described in [40, 37]. In order to simplify the analysis of the transition to the billiard limit, we restrict the class of confining potentials by assuming a power-law growth of V near ∂D (we use the notation BD for assumptions we make on the single particle confined in the bounded domain).

BD1: Power-law growth assumption. *Given any compact subset of the non-singular part of ∂D , in a small neighborhood of this set the potential V is given by*

$$V(q) = \frac{1}{Q(q)^\alpha}, \quad (18)$$

where $\alpha > 0$, and the C^∞ function Q measures the distance to the boundary of D , i.e., $Q(q)|_{q \in \partial D} = 0$ and $\nabla Q(q) \neq 0$; we also choose the sign of Q such that $Q(q) > 0$ inside the domain D .

When we consider N mutually repelling particles moving in the potential field V , their motion is described by the Hamiltonian

$$H = \sum_{n=1, \dots, N} \left[\frac{(p^{(n)})^2}{2} + V(q^{(n)}) \right] + \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W(q^{(n)} - q^{(m)}), \quad (19)$$

where $(q^{(n)}, p^{(n)})$ are coordinates and momenta of the n -th particle and W is a repelling potential, as in Definition 2.1 (note that we do not assume that the interaction potential is small here).

We consider a limit of large energy per particle, namely, we study the behavior at the energy level $H((q^{(1)}, p^{(1)}), \dots, (q^{(N)}, p^{(N)})) = Nh$ for large h . We scale the momenta $p^{(n)}$ to $\sqrt{2h}$, so the Hamiltonian transforms to

$$H = \sum_{n=1, \dots, N} \left[\frac{(p^{(n)})^2}{2} + \delta V(q^{(n)}) \right] + \delta \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W(q^{(n)} - q^{(m)}), \quad (20)$$

where $\delta = \frac{1}{2\hbar}$ (the inverse temperature). Now, the goal is to study the behavior on the fixed energy level $H = \frac{N}{2}$ in the limit $\delta \rightarrow 0^+$.

System (20) is similar to system (2). However, here the single-particle system

$$H_0(q, p; \delta) = \frac{p^2}{2} + \delta V(q), \quad (21)$$

depends on δ in a singular way. The formal limit of the potential energy term as $\delta \rightarrow 0^+$ is the billiard potential, which is zero inside D and infinite at the boundary of D . The corresponding dynamical system, the billiard in the domain D [11], is not smooth, so our Theorems 1 and 2 cannot be directly applied. However, the method we used there can be carried over to this case as well, with the help of an enhanced version of our theory of billiard-like potentials [37, 40].

Recall that the billiard dynamics can be viewed as a motion of a particle along straight segments with speed 1, interrupted by jumps in momenta as the particle reflects from the boundary. The jumps are defined by the elastic reflection law, with the angle of incidence equal to the angle of reflection. Equivalently, the dynamics are determined by the billiard map, which records the position and the angle of reflection at impacts. The dynamics of the smooth system at small δ can be quite different from the dynamics of the formal billiard limit. Still, this formal limit provides good approximation for *regular billiard orbits*, which are defined as orbits for which all impact points are bounded away from singularities of the billiard boundary, and *all the impact angles are bounded away from zero* [37, 40, 51].

Thus, let $L^* = \{(q^*(\omega_0 t), p^*(\omega_0 t))\}$ denote a regular periodic orbit of the billiard in D , which hits the billiard boundary at points M^1, \dots, M^{k^*} (we call them *impact points*, to distinguish from multi-particle *collision points*). Let t^1, \dots, t^{k^*} be the impact moments or time, i.e., $M^j = q^*(\omega_0 t^j)$, $j = 1, \dots, k^*$. The functions q^* and p^* are 2π -periodic. As this is a billiard orbit, p^* is a piece-wise constant function of time, with the jumps of p^* happening at $t = t^j \bmod 2\pi$, $j = 1, \dots, k^*$. The energy conservation implies that $\|p^*\|$ stays constant: $\|p^*\| = 1$; thus, the frequency ω_0 is such that $\frac{2\pi}{\omega_0}$ equals the length of L^* . The function q^* is continuous and piece-wise linear, since $\frac{d}{dt}q^* = p^*$ when $t \neq t^j \bmod 2\pi$, $j = 1, \dots, k^*$. The regularity of the orbit L^* means that the boundary of D is smooth at each of the points M^j and the vectors $p^*(\omega_0 t^j \pm 0)$ are not tangent to the boundary of D at M^j , $j = 1, \dots, k^*$. The impact points M^1, \dots, M^{k^*} comprise a periodic orbit of the billiard map B_D : each of them is a fixed point of the *billiard return map* $(B_D)^{k^*}$. Since the impact points are non-singular and the impacts are non-tangent, this map is smooth in a small neighborhood of any of the impact points.

BD2: Elliptic orbit assumption. *The regular billiard periodic orbit L^* is elliptic and*

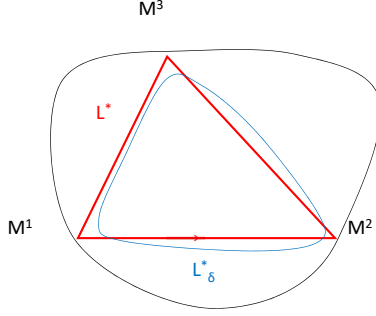


Figure 4: Stable periodic orbit in a convex billiard table (L^*) and a nearby stable periodic orbit (L^*_δ) for the smooth billiard-like potential at high energy ($\delta = \frac{1}{2h}$, see (21)).

KAM-nondegenerate. Namely, the point M^1 is a KAM-nondegenerate elliptic periodic point of the billiard map. This means that two conditions are fulfilled. First, the multipliers ($\exp(\pm i \frac{2\pi}{\omega_0} \omega_1), \dots, \exp(\pm i \frac{2\pi}{\omega_0} \omega_{d-1})$) (the eigenvalues of the derivative of $(B_D)^{k^}$ at the point*

M^1) are non-resonant up to order 4, namely $m_0 \omega_0 + \sum_{j=1}^{d-1} m_j \omega_j \neq 0$ for all integer m_0 and

m_1, \dots, m_{d-1} such that $1 \leq \sum_{j=1}^{d-1} |m_j| \leq 4$. This implies that the Birkhoff normal form for

$(B_D)^{k^}$ in the action-angle coordinates $(I, \Phi) \in \mathbb{R}^{d-1} \times \mathbb{T}^{d-1}$ near M^1 is given by $(I, \Phi) \mapsto (\bar{I}, \bar{\Phi})$, where*

$$\bar{I} = I + o(I), \quad \bar{\Phi} = \Phi + \frac{2\pi}{\omega_0} \hat{\omega} + \Omega I + o(I), \quad (22)$$

with constant $\hat{\omega} = (\omega_1, \dots, \omega_{d-1})$ and $\omega_0 = \frac{2\pi}{|L^|}$, where $|L^*|$ is the length of L^* . The second KAM-nondegeneracy condition (the twist condition) is*

$$\det(\Omega) \neq 0. \quad (23)$$

The existence of a periodic orbit satisfying this assumption holds true for an open set of billiards; for convex billiards in the plane this assumption is also open and dense [14].

The billiard return map $(B_D)^{k^*}$ is smoothly conjugate (by the billiard flow) to the return map to any small cross-section to L^* chosen in the interior of D . As shown in [51, 37], in the limit $\delta \rightarrow +0$, the return map of the smooth flow of (21) on such cross-section tends, in C^∞ , to the return map of the billiard flow. Thus, up to a change of coordinates, the billiard

return map describes the limit of the smooth dynamics defined by (21). In particular, with Assumptions BD1 and BD2, the single-particle system (21) has, for all sufficiently small δ , a KAM-nondegenerate elliptic periodic orbit L_δ^* in the energy level $H_0 = \frac{1}{2}$, which is close to L^* .

Now, we consider the product of N billiard flows in D , with the invariant set as in (3):

$$q^{(n)} = q^*(\omega_0 t + \theta^{(n)}), \quad p^{(n)} = p^*(\omega_0 t + \theta^{(n)}), \quad n = 1, \dots, N. \quad (24)$$

The projection of this invariant set to the Nd -dimensional configuration space is an N -dimensional torus (continuous, but only piece-wise smooth). The averaged potential on this torus is defined as in (10):

$$U(\theta^{(1)}, \dots, \theta^{(N)}) = \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} \frac{1}{T} \int_0^T W(q^*(\omega_0 t + \theta^{(n)}) - q^*(\omega_0 t + \theta^{(m)})) dt. \quad (25)$$

We assume that there exists a line (12) of minima of U :

$$\theta^{(n)} = \theta_{min}^{(n)} + c, \quad n = 1, \dots, N, \quad (26)$$

which satisfies the KAM assumption IP1 of Section 2.2 and the no-collision assumption IP2 of Section 2.4.

The KAM assumption requires a sufficient smoothness of the averaged potential, which does not, a priori, hold for billiard orbits because $q^*(t)$ is not smooth at the impact points. In general, the non-smoothness of the system at $\delta = 0$ can make the averaging procedure invalid and lead to dynamics different from those in the smooth case. However, we show that these issues do not materialize (e.g. we prove the smoothness of the averaged potential, see Lemma 4.7) if the non-interacting particles moving along the same billiard trajectory L^* with the phase shifts $\theta_{min}^{(n)}$ *never hit the billiard boundary simultaneously*:

IP3: Non-simultaneous impacts assumption. *The impacts of $q^*(\omega_0 t + \theta_{min}^{(n)})$ with the billiard boundary do not happen simultaneously, namely, if $\omega_0 t + \theta_{min}^{(n)} = \omega_0 t^j \pmod{2\pi}$ for some j , then $\omega_0 t + \theta_{min}^{(m)} \neq \omega_0 t^k \pmod{2\pi}$ for all k and all $m \neq n$:*

$$\theta_{min}^{(n)} - \theta_{min}^{(m)} \neq \omega_0(t^j - t^k) \pmod{2\pi}. \quad (27)$$

For convenience, we can always assume (by redefining c in (12)) that

$$\sum_{n=1}^N \theta_{min}^{(n)} = 0; \quad (28)$$

by a shift of time, we can also achieve that

$$\theta_{min}^{(n)} \neq \omega_0 t^j, \quad \text{for all } n = 1, \dots, N, \quad j = 1, \dots, k^*. \quad (29)$$

Theorem 3. *Consider N repelling particles that are confined to a region D by a trapping potential satisfying the power-law assumption BD1. Assume that the billiard table D has a regular elliptic periodic orbit L^* which satisfies the elliptic orbit assumption BD2, and that the averaged interaction potential has a minima line satisfying the KAM assumption IP1, the no-collision assumption IP2, and the non-simultaneous impacts assumption IP3. Then, for all sufficiently high values of the energy-per-particle h , the N -particle system (19) has a positive measure set of initial conditions corresponding to quasi-periodic solutions as in Theorem 1, with $\bar{\omega} = \frac{2\pi\sqrt{2h}}{|L^*|}(1 + o_{h \rightarrow \infty}(1))$. In particular, this system is not ergodic for all sufficiently high energies.*

The proof is in Section 4. It is an empirical fact that Hamiltonian systems with low number of degrees of freedom have elliptic periodic orbits easily, unless the system is specially prepared to have a (partially) hyperbolic structure on every energy level. Therefore, a common belief (and a challenging conjecture to prove) is that a generic Hamiltonian system without the uniform partially-hyperbolic structure possesses a non-degenerate elliptic orbit. The billiard counterpart of such claim would be that a generic billiard which is not of the dispersing or defocusing type [53] has a non-degenerate elliptic orbit. Currently, no methods are known for proving such conjecture in any reasonable regularity class. But, once we accept this conjecture for systems with low number of degrees of freedom, Theorem 3 implies that the gas of any number of repelling particles confined in a domain with a sufficiently smooth boundary is *generically non-ergodic* for all sufficiently high temperatures.

2.6 Particles in a rectangular box

The single-particle billiard in a rectangular box has no elliptic periodic orbits: it is an integrable system (with partial oscillatory motions parallel to different coordinate axes independent of each other), so the periodic orbits are parabolic. The orbits of the same period go in several continuous $(d-1)$ -parameter families; the orbits in the same family can be distinguished by the phase differences between the partial oscillations or by the coordinates of

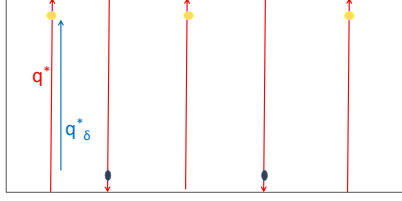


Figure 5: In a rectangular box, synchronized parallel motion by which all particles move vertically and hit the boundary simultaneously is KAM-stable under the conditions of Theorem 5. Less ordered parallel motion, by which all particles have the same period yet each particle turns at a different moment, is KAM-stable under the conditions of Theorem 4.

the impact points. Namely, if the box sizes are (l_1, \dots, l_d) and the conserved kinetic energies of the corresponding partial oscillations are $E_j = \frac{1}{2}p_j^2$, then the frequencies of the partial oscillations are $2\pi\frac{|p_j|}{l_j}$, and the single-particle motion is periodic if and only if the ratio of each two of these frequencies is a rational number. Thus, we have a discrete set of possible choices of partial energies, for which the motion with any initial point (q_1, \dots, q_d) in the box is periodic with the same period (for any choice of the signs of p_j).

While similar computations can be performed for any of these families, we choose the simplest one, where all the particles move strictly along one of the coordinate axes, i.e., the family is given by the equation $p_1 = \dots = p_{d-1} = 0$. We call such oscillations vertical; the particle moves up for a half of the period and it moves down for the other half. When the energy is fixed, different periodic orbits in this family are distinguished by the values of the “horizontal” coordinates (q_1, \dots, q_{d-1}) , which do not change with time. In the same spirit as before, one can place any number of non-interacting particles on this family (each particle with the same kinetic energy, but on its own path, i.e., with different values of the horizontal coordinates). The difference with the previous cases is that we now allow the particles to spread over a continuous family and not over just one orbit. If the particle energy is high enough, switching the repulsion between the particles on makes only a small perturbation of the fast vertical motion. We show below that the slow evolution of the horizontal degrees of freedom and the differences between the phases of the vertical oscillations are governed, to the main order, by the averaged potential; its non-degenerate minima correspond to elliptic orbits of the multi-particle system.

Note that we have only one fast degree of freedom for the entire multi-particle system in this setting (the sum of the phases of the vertical oscillations). This makes the averaging procedure simpler than in the previous cases. However, the non-simultaneous impacts assumption IP3, which is crucial for justification of the averaging in Theorem 3, can be violated for the family of vertically oscillating particles for an open set of repelling potentials (see below). We therefore develop a different approach for the case of simultaneous impacts.

Let us describe the assumptions we impose on the system in the box.

Box1: Separability assumption. *The single-particle Hamiltonian is given by*

$$H_0(q, p) = \sum_{i=1}^d \left(\frac{p_i^2}{2} + V_i(q_i) \right), \quad (30)$$

with

$$V_i(q_i) = \frac{1}{Q_i(q_i)^\alpha}, \quad \alpha > 0 \quad (31)$$

where the C^∞ function Q_i measures the distance to the box boundary in the i -th coordinate direction, i.e., $Q_i(0) = 0$, $Q_i(l_i) = 0$, $Q_i'(0) > 0$, $Q_i'(l_i) < 0$, and $Q_i(q_i) > 0$ for $q_i \in (0, l_i)$. Finally assume that the potential is symmetric⁵ in the vertical direction: $Q_d(l_d - q_d) = Q_d(q_d)$.

Thus, the N -particle Hamiltonian has the form

$$H = \sum_{n=1, \dots, N} \sum_{i=1}^d \left[\frac{(p_i^{(n)})^2}{2} + V_i(q_i^{(n)}) \right] + \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W(q^{(n)} - q^{(m)}), \quad (32)$$

where W is a repelling potential, C^∞ for $\|q^{(n)} - q^{(m)}\| > \rho$ (see Definition 2.1).

We consider the limit of large energy per particle, and look for motions which are fast only in the last coordinate. Namely, we study the behavior at the energy level $H = Nh$ for a fixed N and large h where most of the particles' energy is at the vertical motion. We scale the vertical momenta $p_d^{(n)}$ by $\sqrt{2h}$, and the Hamiltonian transforms to

$$H = \sum_{n=1, \dots, N} \left[\frac{(p_d^{(n)})^2}{2} + \delta V_d(q_d^{(n)}) \right] + \delta \sum_{n=1, \dots, N} \sum_{i=1}^{d-1} \left[\frac{(p_i^{(n)})^2}{2} + V_i(q_i^{(n)}) \right] + \delta \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W(q^{(n)} - q^{(m)}), \quad (33)$$

where $\delta = \frac{1}{2h}$; as in section 2.5, we study the behavior on the fixed energy level $H = \frac{N}{2}$ in the limit $\delta \rightarrow +0$.

In the limit $\delta = 0$, the Hamiltonian describes N independent vertical, constant speed,

⁵This symmetry assumption appears to be non-essential. It is not used at all in Theorem 4. We include it here as it is natural and makes some notations and computations in the proof of Theorem 5 simpler.

saw-tooth motions. Setting all the particles to have the same speed $|p_d^{(n)}| = 1$ in the limit $\delta = 0$, and choosing the vertical size of the box $l_d = \pi$, we obtain the limiting family of solutions in the form $q_i^{(n)}(t) = \text{constant}$ for $i = 1, \dots, d-1$ and $q_d^{(n)}(t) = q^*(t + \theta^{(n)})$ where q^* is the 2π -periodic saw-tooth function:

$$q^*(t) = \begin{cases} t, & \text{for } t \in [0, \pi], \\ 2\pi - t, & \text{for } t \in [\pi, 2\pi]. \end{cases} \quad (34)$$

Denote the horizontal coordinates of the n -th particle by $\xi^{(n)}$ (so $\xi^{(n)} = (\xi_1^{(n)}, \dots, \xi_{d-1}^{(n)}) = (q_1^{(n)}, \dots, q_{d-1}^{(n)})$). Define the averaged potential,

$$U(\theta, \xi) = \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}) + \sum_{n=1, \dots, N} \sum_{i=1}^{d-1} V_i(\xi_i^{(n)}) \quad (35)$$

where

$$W_{avg} = \frac{1}{2\pi} \int_0^{2\pi} W(q^*(s + \theta^{(n)}) - q^*(s + \theta^{(m)}), \xi^{(n)} - \xi^{(m)}) ds, \quad (36)$$

and $\theta = (\theta^{(1)}, \dots, \theta^{(N)})$, $\xi = (\xi_1^{(1)}, \dots, \xi_{d-1}^{(1)}, \dots, \xi_1^{(N)}, \dots, \xi_{d-1}^{(N)})$. Let $\|\xi^{(n)} - \xi^{(m)}\| > \rho$ for every $n \neq m$. We establish in Section 5 (see Lemma 5.2) that the averaged potential is C^∞ -smooth if $\theta^{(n)} \neq \theta^{(m)} \pmod{\pi}$ for every $n \neq m$. Moreover, we also show that under the parity assumption Box4 below, the averaged potential is, along with all its derivatives with respect to ξ , at least C^2 -smooth function of θ even if $\theta^{(n)} = \theta^{(m)} \pmod{\pi}$ for some, or all, $n \neq m$.

Since V_i and W_{avg} are bounded from below, the averaged potential $U(\theta, \xi)$ must have a minimum line

$$\theta^{(n)} = \theta_{min}^{(n)} + c, \quad \xi^{(n)} = \xi_{min}^{(n)}, \quad (37)$$

where c is an arbitrary constant. Like in Section 2.2, we can introduce coordinates (φ, ψ, ξ) in a small neighborhood of this line such that $\varphi = \frac{1}{N}(\theta^{(1)} + \dots + \theta^{(N)})$ and the coordinates $\psi = (\psi_1, \dots, \psi_{N-1})$ are linear combinations of the phase differences $(\theta^{(n)} - \theta_{min}^{(n)}) - (\theta^{(m)} - \theta_{min}^{(m)})$. Since the averaged potential U depends only on the differences of the phases, we obtain that it is independent of φ , so, as in (13), we set

$$U(\theta, \xi) = \hat{U}(\psi, \xi).$$

Box2: Non-degenerate minimum assumption. *The minimum of the averaged potential \hat{U} corresponds to $\|\xi_{min}^{(n)} - \xi_{min}^{(m)}\| > \rho$ for all $n \neq m$. The Hessian matrix of \hat{U} at the minimum*

$\psi = 0, \xi = \xi_{min}$ is non-degenerate, and all its eigenvalues are simple.

The condition $\|\xi_{min}^{(n)} - \xi_{min}^{(m)}\| > \rho$ means that the particles stay away from each other, each on its own path. This assumption is fulfilled automatically when, for example, W satisfies (16). Indeed, then, by Lemma 2.2, the minimum of the averaged potential cannot correspond to collisions, yet, two particles on the same vertical path collide unavoidably.

The non-degeneracy of the Hessian is a generic condition, implying that $(\psi = 0, \xi = \xi_{min})$ corresponds to the elliptic equilibrium of the Hamiltonian

$$H = \frac{1}{2N} J^2 + \frac{p_\xi^2}{2} + \hat{U}(\psi, \xi), \quad (38)$$

where J and p_ξ denote the conjugate momenta to ψ and ξ , respectively.

Let us first consider the case of *non-simultaneous impacts* motion at which $\theta_{min}^{(n)} \neq \theta_{min}^{(m)} \pmod{\pi}$ for all $n \neq m$ (a particle impacts to the boundary happen exactly at each half-period, so this condition, obviously, means that no two particles hit the boundary simultaneously). By Lemma 5.1 the averaged potential near the minimum is C^∞ -smooth, so we impose the following genericity condition (which involves the Taylor expansion up to order 4).

Box3: KAM assumption (the case of non-simultaneous impacts). *The local minimum of \hat{U} at $(\psi = 0, \xi = \xi_{min})$ is KAM-non-degenerate: the corresponding elliptic equilibrium of the Hamiltonian system (38) has no resonances up to order 4 and its Birkhoff normal form satisfies the twist condition.*

Theorem 4. *Consider N repelling particles that are confined to a box by a trapping potential satisfying the separability assumption Box1. Let the averaged interaction potential have a minima line, corresponding to non-simultaneous impacts and satisfying the nondegeneracy assumptions Box2 and Box3. Then, for all sufficiently high values h of the energy per particle, the N -particle system (32) has a non-degenerate elliptic periodic orbit accompanied by a positive measure set of quasi-periodic solutions. In particular, for this set of initial conditions, each particle stays bounded away from all other particles for all time, so the system is not ergodic.*

The assumption of non-simultaneous impacts is generic when the interaction potential has no special symmetries. However, for the most natural class of potentials which depend only on the Euclidian distance between particles, there is an inherit symmetry which can lock the impacts to become simultaneous. Such potentials satisfy

Box4: Parity assumption. *The repelling interaction potential W is even in $q_i^{(n)} - q_i^{(m)}$ for each $i = 1, \dots, d$.*

In this case, the average potential U is an even function of $(\theta^{(n)} - \theta^{(m)})$, for any pair of

n and m . It is also 2π -periodic in $(\theta^{(n)} - \theta^{(m)})$. We conclude that if Box4 is satisfied, then

$$U(\theta^{(n)} - \theta^{(m)}, \cdot) = U(\theta^{(m)} - \theta^{(n)}, \cdot) \text{ and } U(\theta^{(n)} - \theta^{(m)} - \pi, \cdot) = U(\pi - (\theta^{(n)} - \theta^{(m)}), \cdot). \quad (39)$$

It follows that the first derivative of U with respect to θ vanishes when $\theta^{(n)} = \theta^{(m)} \pmod{\pi}$ for all m and n . Therefore, when the parity assumption holds, there can exist minima of U for which *all the particles hit the boundary walls simultaneously* (each half-period, some particles hit $q_d = 0$ while, at the same time, the others hit $q_d = \pi$); moreover, this *simultaneous impacts* property can persist for small perturbations of the potential W within the class of potentials satisfying Assumption Box4. The non-simultaneous impacts assumption, which is crucial for the averaged procedure we use in Theorems 3 and 4, does not hold for such minima. We, therefore, consider this case separately.

First note that the averaged potential U in the simultaneous impact case is not, in general, C^3 -smooth with respect to θ (see Lemma 5.2). Hence, there cannot be a direct analogue of the KAM-nondegeneracy assumption Box3, which involves derivatives of U up to order 4 (recall that the twist condition on quadratic terms in the action variables corresponds to algebraic relation between the Taylor coefficients of the Hamiltonian up to order 4 [3]). Instead of assumption Box3 which is formulated in terms of the Hamiltonian system (38), we formulate KAM nondegeneracy assumption in terms of the two auxiliary Hamiltonians:

$$H_0^\xi(p_\xi, \xi) = \frac{p_\xi^2}{2} + U(\theta_{min}, \xi), \quad (40)$$

and

$$H_0^\theta(p_d, \theta) = \sum_{n=1, \dots, N} \frac{(p_d^{(n)})^2}{2} + \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} \gamma_{nm} (\theta^{(n)} - \theta^{(m)} - \vartheta_{nm})^2 - \beta_{nm} (\theta^{(n)} - \theta^{(m)} - \vartheta_{nm})^4, \quad (41)$$

where $\vartheta_{nm} = \theta_{min}^{(n)} - \theta_{min}^{(m)} \pmod{2\pi}$ is either 0 or π , and

$$\gamma_{nm} = \begin{cases} \frac{\partial^2 W}{\partial q_d^2}(0, \xi_{min}^{(n)} - \xi_{min}^{(m)}) & \text{for } \vartheta_{nm} = 0, \\ -\frac{1}{\pi} \frac{\partial W}{\partial q_d}(\pi, \xi_{min}^{(n)} - \xi_{min}^{(m)}) & \text{for } \vartheta_{nm} = \pi, \end{cases}$$

$$\beta_{nm} = \begin{cases} -\frac{\partial^2 W}{\partial q_d^2}(0, \xi_{min}^{(n)} - \xi_{min}^{(m)}) & \text{for } \vartheta_{nm} = 0, \\ \frac{\partial^2 W}{\partial q_d^2}(\pi, \xi_{min}^{(n)} - \xi_{min}^{(m)}) & \text{for } \vartheta_{nm} = \pi. \end{cases}$$

The quadratic part of the potential in H_0^θ coincides with the quadratic term of the Taylor

expansion of $U(\theta, \xi_{min})$ at $\theta = \theta_{min}$ (see Lemma 5.3). Due to the simultaneous impacts property, by redefining the constant c in (37), if necessary, we can always make $\theta_{min}^{(n)} = 0 \pmod{\pi}$. Then, by (39), the averaged potential U is an even function of $(\theta - \theta_{min})$, so the Hessian of U is block-diagonal: the derivatives $\frac{\partial^2 U}{\partial \theta \partial \xi}$ vanish at the minimum with simultaneous impacts. Therefore, the non-degenerate minimum assumption Box2 implies that both $U(\theta_{min}, \xi)$ and $\hat{U}(\psi, \xi_{min})$ have non-degenerate minima (at $\xi = \xi_{min}$ and at $\psi = 0$, respectively). We conclude that, under Assumption Box2, Hamiltonian (40) has an elliptic equilibrium at $\xi = \xi_{min}$, $p_\xi = 0$, and Hamiltonian (41) has a family of elliptic periodic orbits $\theta^{(n)} = \theta_{min}^{(n)} + \omega t$ ($n = 1, \dots, N$), $|p_d^{(1)}| = \dots = |p_d^{(N)}| = \omega = const$. This is a relative equilibrium, i.e., it becomes an equilibrium when we go to translation invariant (J, ψ) -coordinates, like in Section 2.2. Denote the reduced Hamiltonian of (41) by $H^{\theta,0}(J, \psi)$.

Box5: KAM assumption (the case of simultaneous impacts). *Let a nondegenerate minimum line of U satisfy $\theta_{min}^{(n)} = \theta_{min}^{(m)} \pmod{\pi}$ for all m and n . Assume that at the local minimum of \hat{U} at $(\psi = 0, \xi = \xi_{min})$ the frequencies of small oscillations have no resonances up to order 4. Furthermore, assume that the elliptic equilibrium of H_0^ξ at $p_\xi = 0, \xi = \xi_{min}$ and the elliptic equilibrium of $H^{\theta,0}$ at $(J, \psi) = 0$ are KAM-non-degenerate, i.e. their Birkhoff normal forms satisfy the twist condition.*

Theorem 5. *Consider N repelling particles that are confined to a box by a trapping potential satisfying the separability assumption Box1 with $\alpha > 6$. Assume the parity assumption Box4 holds. Assume the averaged interaction potential U has a minima line with simultaneous impacts and let it satisfy the non-degeneracy assumptions of Box2 and Box5. Then, for all sufficiently high values of the energy per particle, the N -particle system (32) has a non-degenerate elliptic periodic orbit accompanied by a positive measure set of quasi-periodic solutions. In particular, for this set of initial conditions, each particle stays bounded away from all other particles for all time, so the system is not ergodic.*

Thus, the gas of any number of highly-energetic repelling particles confined to a rectangular box by a sufficiently steep potential is, generically, non-ergodic.

3 Smooth N-particle systems.

We present here the proof of Theorem 1. Its outline is as follows. First, we consider the N -particle system (2) near $\mathbf{L}^*(\theta_{min})$, the choreographic periodic orbit of the uncoupled system, and scale the action coordinates by $\delta^{1/2}$. Then we average, i.e. we make a change of coordinates, after which the angle-dependent terms become $O(\delta^{3/4})$, see Lemma 3.1. Next, we show that near a line of minima, the Poincaré map of the system is $O(\delta^{3/4})$ -close to the

flow map of the Hamiltonian system of a specific form (55) (Lemma 3.2). We need to take $O(1/\delta^{1/2})$ iterates of the return map. In Lemma 3.3, we establish that taking $O(1/\delta^{1/2})$ iterates of any map which is $O(\delta^{3/4})$ close to the return map of system (55) leads to a map which is $O(\delta^{1/4})$ close to the time-1 map for the Hamiltonian (58) near the origin. In Lemma 3.4, we prove that the Hamiltonian (58) admits a positive measure set of KAM tori. By Lemma 3.3, one infers the results of Theorem 1 from the KAM theorem.

Proof: Recall that $(q^{(n)}, p^{(n)})$ denotes the coordinates and conjugate momenta of the n -th particle, $n = 1, \dots, N$. We apply the same symplectic transformation that brings the single-particle system to the normal form (6) to each of the N particles, namely, we let $(q^{(n)}, p^{(n)}) = (\hat{q}(I_0^{(n)}, \theta^{(n)}, z^{(n)}), \hat{p}(I_0^{(n)}, \theta^{(n)}, z^{(n)}))$. As for the single particle case, this change of coordinates is smooth and preserves the standard symplectic form. In these coordinates the N -particle system (2) takes the form:

$$H = \sum_{n=1}^N [\omega I^{(n)} + \frac{1}{2} I^{(n)} A I^{(n)} + g(I_0^{(n)}, \theta^{(n)}, z^{(n)})] + \delta \sum_{n \neq m} W(\hat{q}(I_0^{(n)}, \theta^{(n)}, z^{(n)}) - \hat{q}(I_0^{(m)}, \theta^{(m)}, z^{(m)})), \quad (42)$$

where we denote $I^{(n)} = (I_0^{(n)}, I_1^{(n)} = \frac{(z_1^{(n)})^2}{2}, \dots, I_{d-1}^{(n)} = \frac{(z_{d-1}^{(n)})^2}{2})$. We will look at the motion of the particles near the periodic trajectory L^* . Namely, we will write

$$\hat{q}(I_0^{(n)}, \theta^{(n)}, z^{(n)}) = q^*(\theta^{(n)}) + O(|I_0^{(n)}| + \|z^{(n)}\|). \quad (43)$$

Since the pairwise interaction potential W is smooth,

$$W(\hat{q}(I_0^{(n)}, \theta^{(n)}, z^{(n)}) - \hat{q}(I_0^{(m)}, \theta^{(m)}, z^{(m)})) = W(q^*(\theta^{(n)}) - q^*(\theta^{(m)})) + O(|I_0^{(n)}| + \|z^{(n)}\| + |I_0^{(m)}| + \|z^{(m)}\|). \quad (44)$$

We restrict our attention to the region of the phase space where the actions $I^{(n)}$ are of order $\delta^{1/2}$. For that, we scale the variables $z^{(n)}$ to $\delta^{1/4}$ and the variables $I_0^{(n)}$ to $\delta^{1/2}$, i.e., we make a replacement $z^{(n)} \rightarrow \delta^{1/4} z^{(n)}$, $I_0^{(n)} \rightarrow \delta^{1/2} I_0^{(n)}$. With this scaling, we have

$$W = W(q^*(\theta^{(n)}) - q^*(\theta^{(m)})) + O(\delta^{1/4}),$$

(see (44)) and so, by (6), the remainder terms in (42), g , satisfy

$$g(I_0^{(n)}, \theta^{(n)}, z^{(n)}) = O(\delta^{5/4}).$$

The motion in the rescaled variables is described by the rescaled Hamiltonian

$$\begin{aligned} H_{scal} &= \delta^{-1/2} H\left((\delta^{1/2} I_0^{(n)}, \theta^{(n)}, \delta^{1/4} z^{(n)})_{n=1, \dots, N}\right) \\ &= \sum_{n=1}^N \omega I^{(n)} + \delta^{1/2} \sum_{n=1}^N \frac{1}{2} I^{(n)} A I^{(n)} + \delta^{1/2} \sum_{n \neq m} W(q^*(\theta^{(n)}) - q^*(\theta^{(m)})) + O(\delta^{3/4}), \end{aligned} \quad (45)$$

We now average the Hamiltonian with respect to the motion along the periodic orbit.

Lemma 3.1. *There exists a smooth symplectic change of coordinates which brings the Hamiltonian (45) to the form*

$$H = \sum_{n=1}^N \omega I^{(n)} + \delta^{1/2} \sum_{n=1}^N \frac{1}{2} I^{(n)} A I^{(n)} + \delta^{1/2} U(\theta^{(1)}, \dots, \theta^{(N)}) + O(\delta^{3/4}), \quad (46)$$

where the averaged potential U is given by (10).

Proof. Recall that the (non-averaged) interaction potential

$$\tilde{W}(\theta^{(1)}, \dots, \theta^{(N)}) = \sum_{n \neq m} W(q^*(\theta^{(n)}) - q^*(\theta^{(m)}))$$

is 2π -periodic in each of the variables $\theta^{(n)}$, $n = 1, \dots, N$. Therefore, we can write its Fourier expansion:

$$\tilde{W}(\theta^{(1)}, \dots, \theta^{(N)}) = \sum_{n \neq m} \sum_{(k_1, k_2) \in \mathbb{Z}^2} w_{k_1, k_2, n, m} e^{i(k_1 \theta^{(n)} + k_2 \theta^{(m)})}. \quad (47)$$

The function \tilde{W} is of class C^∞ , so the Fourier coefficients w decay fast as $k_{1,2}$ grow. In particular, the series

$$\Psi(\theta^{(1)}, \dots, \theta^{(N)}) = -i \sum_{n \neq m} \sum_{k_1 + k_2 \neq 0} \frac{w_{k_1, k_2, n, m}}{\omega_0(k_1 + k_2)} e^{i(k_1 \theta^{(n)} + k_2 \theta^{(m)})} \quad (48)$$

is absolutely convergent, and the sum is a C^∞ function of $\theta^{(1)}, \dots, \theta^{(N)}$. By construction,

$$\omega_0(\partial_{\theta^{(1)}} \Psi + \dots + \partial_{\theta^{(N)}} \Psi) = \tilde{W} - U, \quad (49)$$

where

$$U = \sum_{n \neq m} \sum_{k_1 + k_2 = 0} w_{k_1, k_2, n, m} e^{i(k_1 \theta^{(n)} + k_2 \theta^{(m)})} = \sum_{n \neq m} \sum_{k \in \mathbb{Z}} w_{k, -k, n, m} e^{ik(\theta^{(n)} - \theta^{(m)})}.$$

Substituting (47) in (11) and integrating over time shows that the above U is indeed the averaged potential given by (10).

Now, we perform a symplectic coordinate change

$$I_0^{(n)} \rightarrow I_0^{(n)} - \delta^{1/2} \partial_{\theta^{(n)}} \Psi, \quad n = 1, \dots, N \quad (50)$$

(all other variables remain unchanged). The first term in the Hamiltonian (45),

$$\sum_{n=1}^N \omega I^{(n)} = \omega_0 (I_0^{(1)} + \dots + I_0^{(N)}) + \sum_{n=1}^N \sum_{j=1}^{d-1} \omega_j I_j^{(n)},$$

after substituting (50), produces, by (49), the additional term $-\delta^{1/2}(\tilde{W} - U)$. Substituting (50) in the $O(\delta^{1/2})$ terms leads only to $O(\delta)$ corrections. Hence, the Hamiltonian takes the required form (46). \square

Now, as explained in Section 2, we utilize the translation symmetry of the averaged potential $U(\theta^{(1)}, \dots, \theta^{(N)})$ and introduce the collective phase $\varphi = \frac{1}{N}(\theta^{(1)} + \dots + \theta^{(N)})$ and coordinates $\psi = (\psi_1, \dots, \psi_{N-1})$ which measure small deviations of the longitudinal motion from the line (12) of minima defined by θ_{min} . The precise definition of the coordinates ψ is as follows. We choose an $(N \times N)$ orthogonal matrix R such that its first row is $\frac{1}{\sqrt{N}}(1, \dots, 1)$, and define

$$(\varphi, \psi)^\top = \frac{1}{\sqrt{N}} R \begin{pmatrix} \theta^{(1)} - \theta_{min}^{(1)} \\ \vdots \\ \theta^{(N)} - \theta_{min}^{(N)} \end{pmatrix}. \quad (51)$$

By construction, the second and further rows of R are all orthogonal to $(1, \dots, 1)$, which implies that $\psi = 0$ for every point of the line (12). Notice that while φ is an angle (the $O(\delta^{3/4})$ term in (46) is periodic in φ), the variables ψ_n correspond to small deviations from the minimum and we do not define them globally (so they are not angular variables). In the new coordinates the averaged potential U is independent of φ , and is given by $\hat{U}(\psi)$ of (13). Next, we define conjugate momenta (P, J_1, \dots, J_{N-1}) corresponding to the variables (φ, ψ) :

$$(P, J)^\top = \sqrt{N} R \begin{pmatrix} I_0^{(1)} \\ \vdots \\ I_0^{(N)} \end{pmatrix}. \quad (52)$$

In particular,

$$P = \sum_{n=1}^N I_0^{(n)}.$$

Since the first row of the orthogonal matrix R is $\frac{1}{\sqrt{N}}(1, \dots, 1)$, it follows that

$$I_0^{(n)} = \frac{1}{\sqrt{N}} \left(\frac{P}{\sqrt{N}} + \sum_{m=1}^{N-1} R_{m+1,n} J_m \right).$$

This transformation $(\theta_0^{(1)}, \dots, \theta_0^{(N)}, I_0^{(1)}, \dots, I_0^{(N)}) \mapsto (\varphi, \psi, P, J)$ is defined by the generating function $\sqrt{N}(\varphi, \psi)R(I_0^{(1)}, \dots, I_0^{(N)})^\top$, so it is symplectic.

Therefore, we can perform this transformation in the Hamiltonian function (46) directly. As a result, we obtain the new Hamiltonian (recall that the matrix A in (46) is given by (7)):

$$\begin{aligned} H &= \omega_0 P + \sum_{n=1}^N \sum_{j=1}^{d-1} \omega_j I_j^{(n)} + \\ &+ \delta^{1/2} \left(a \frac{P^2 + J^2}{2N} + \sum_{n=1}^N \frac{1}{\sqrt{N}} \left(\frac{P}{\sqrt{N}} + \sum_{m=1}^{N-1} R_{m+1,n} J_m \right) \sum_{j=1}^{d-1} b_j I_j^{(n)} + \frac{1}{2} \sum_{n=1}^N \sum_{i,j=1}^{d-1} \hat{a}_{ij} I_j^{(n)} I_i^{(n)} \right) + \\ &+ \delta^{1/2} \hat{U}(\psi) + O(\delta^{3/4}). \end{aligned} \tag{53}$$

Denoting $\hat{\omega} = (\omega_1, \dots, \omega_{d-1})$, $\hat{I}^{(n)} = (I_1^{(n)}, \dots, I_{d-1}^{(n)})^\top$ (recall that $I_j^{(n)} = \frac{(z_j^{(n)})^2}{2} = \frac{(x_j^{(n)})^2 + (y_j^{(n)})^2}{2}$), and $R_n = (R_{2,n}, \dots, R_{N,n})^\top$, the Hamiltonian recasts as

$$\begin{aligned} &H(P, J, \varphi, \psi, \{z_j^{(n)}\}_{j=1, \dots, d-1, n=1, \dots, N}) \\ &= \omega_0 P + \sum_{n=1}^N \hat{\omega} \hat{I}^{(n)} + \delta^{1/2} \left(a \frac{P^2 + J^2}{2N} + \frac{1}{\sqrt{N}} \sum_{n=1}^N \left(\frac{P}{\sqrt{N}} + R_n^\top J \right) b \hat{I}^{(n)} + \sum_{n=1}^N \frac{1}{2} \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} \right) + \\ &+ \delta^{1/2} \hat{U}(\psi) + O(\delta^{3/4}). \end{aligned} \tag{54}$$

Note that $\dot{\varphi} = \partial_P H = \omega_0 + O(\delta^{1/2}) > 0$ in this system. Therefore, the Poincaré return map from the hypersurface $\varphi = 0$ to $\varphi = 2\pi$ (i.e., to itself) is well-defined.

Lemma 3.2. *The Poincaré return map for system (54) restricted to the energy level $H = h$ is $O(\delta^{3/4})$ -close to the time- $\frac{2\pi}{\omega_0}$ map of the system*

$$\begin{aligned} \dot{j} &= -\delta^{1/2} \partial_\psi \hat{U}(\psi), & \dot{\psi} &= \delta^{1/2} \left(\frac{a}{N} J + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b \hat{I}^{(n)}) R_n \right), \\ \dot{x}_j^{(n)} &= \Omega_j^{(n)} y_j^{(n)}, & \dot{y}_j^{(n)} &= -\Omega_j^{(n)} x_j^{(n)} \quad (j = 1, \dots, d-1; n = 1, \dots, N), \end{aligned} \tag{55}$$

where

$$\begin{aligned} \Omega_j^{(n)} &= \frac{h}{\omega_0 N} (b_j - \frac{a}{\omega_0} \omega_j) + \frac{1}{\omega_0 \sqrt{N}} (R_n^\top J) b_j + \hat{A}_j \hat{I}^{(n)} - \\ &- \frac{1}{\omega_0 N} \sum_{l=1}^{d-1} \sum_{m=1}^N (\omega_j b_l + b_j \omega_l - \frac{a}{\omega_0} \omega_l \omega_j) I_l^{(m)}; \end{aligned} \tag{56}$$

here b_j stands for the j -th element of the vector b , and \hat{A}_j for the j -th row of the matrix \hat{A} .

Proof. The system of differential equations defined by Hamiltonian (54) is

$$\begin{aligned}\dot{P} &= O(\delta^{3/4}), \\ \dot{\varphi} &= \omega_0 + \delta^{1/2} \left(\frac{a}{N} P + \frac{1}{N} \sum_{n=1}^N b \hat{I}^{(n)} \right) + O(\delta^{3/4}), \\ \dot{J} &= -\delta^{1/2} \partial_\psi \hat{U}(\psi) + O(\delta^{3/4}), \\ \dot{\psi} &= \delta^{1/2} \left(\frac{a}{N} J + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b \hat{I}^{(n)}) R_n \right) + O(\delta^{3/4}),\end{aligned}\tag{57}$$

$$\begin{aligned}\dot{x}_j^{(n)} &= \partial_{y_j^{(n)}} H = \tilde{\Omega}_j^{(n)} y_j^{(n)} + O(\delta^{3/4}), \\ \dot{y}_j^{(n)} &= -\partial_{x_j^{(n)}} H = -\tilde{\Omega}_j^{(n)} x_j^{(n)} + O(\delta^{3/4}) \quad (j = 1, \dots, d-1; n = 1, \dots, N),\end{aligned}$$

where $\tilde{\Omega}_j^{(n)} = \frac{\partial H}{\partial I_j^{(n)}} = \omega_j + \delta^{1/2} \left(\frac{1}{\sqrt{N}} \left(\frac{P}{\sqrt{N}} + R_n^\top J \right) b_j + \hat{A}_j \hat{I}^{(n)} \right)$. Applying the inverse function theorem to (54), we can express P as a function of all other variables on the energy level $H = h$

$$P = \frac{1}{\omega_0} \left(h - \hat{\omega} \sum_{m=1}^N \hat{I}^{(m)} \right) + O(\delta^{1/2}).$$

We substitute this expression into (57) and choose φ/ω_0 as the new time variable (i.e., we divide \dot{J} , $\dot{\psi}$, $\dot{x}_j^{(n)}$, and $\dot{y}_j^{(n)}$ to $\dot{\varphi}/\omega_0$). One can see that the result is $O(\delta^{3/4})$ -close to system (55). Since the sought Poincaré map is the time- $\frac{2\pi}{\omega_0}$ map in the new time, we immediately obtain the lemma. \square

Lemma 3.3. *Let $K = \lfloor \frac{\omega_0}{2\pi\delta^{1/2}} \rfloor$ and $\nu = (\nu_1, \dots, \nu_{d-1})$, where $\nu_j = 2\pi \left\{ K \frac{\omega_j}{\omega_0} \right\}$, $j = 1, \dots, d-1$. Then, for all small h , the K -th iteration of any map which is $O(\delta^{3/4})$ -close to the time- $\frac{2\pi}{\omega_0}$ map of system (55), is $O(\delta^{1/4})$ -close to the time-1 map of the flow defined by the Hamiltonian*

$$\begin{aligned}\bar{H} &= \frac{a}{2N} J^2 + \hat{U}(\psi) + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b \hat{I}^{(n)}) (R_n^\top J) + \left(\nu + \frac{h}{\omega_0 N} \left(b - \frac{a}{\omega_0} \hat{\omega} \right) \right) \sum_{n=1}^N \hat{I}^{(n)} + \\ &+ \frac{1}{2} \sum_{n=1}^N \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \hat{I}^{(n)} S \hat{I}^{(m)},\end{aligned}\tag{58}$$

where the symmetric matrix S is given by

$$S = \frac{1}{\omega_0 N} (\hat{\omega}^\top b + b^\top \hat{\omega} - \frac{a}{\omega_0} \hat{\omega}^\top \hat{\omega}),\tag{59}$$

namely

$$S_{jl} = \frac{1}{\omega_0 N} (\omega_j b_l + \omega_l b_j - \frac{a}{\omega_0} \omega_l \omega_j), \quad j, l = 1, \dots, d-1.$$

Proof. Denote $(x_j^{(n)}, y_j^{(n)}) = \sqrt{2I_j^{(n)}} (\cos \phi_j^{(n)}, \sin \phi_j^{(n)})$. For system (55), the actions $I_j^{(n)} = \frac{(x_j^{(n)})^2 + (y_j^{(n)})^2}{2}$ are constants of motion. So, for this system, the time- $\frac{2\pi}{\omega_0}$ map $(J, \psi, I_j^{(n)}, \phi_j^{(n)}) \mapsto (\bar{J}, \bar{\psi}, \bar{I}_j^{(n)}, \bar{\phi}_j^{(n)})$ ($j = 1, \dots, d-1; n = 1, \dots, N$) is given by

$$\begin{aligned} \bar{J} &= J - \frac{2\pi}{\omega_0} \delta^{1/2} \partial_\psi \hat{U}(\psi) + O(\delta), \\ \bar{\psi} &= \psi + \frac{2\pi}{\omega_0} \delta^{1/2} \left(\frac{a}{N} J + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b \hat{I}^{(n)}) R_n \right) + O(\delta), \\ \bar{I}_j^{(n)} &= I_j^{(n)}, \\ \bar{\phi}_j^{(n)} &= \phi_j^{(n)} + \frac{2\pi}{\omega_0} \omega_j + \frac{2\pi}{\omega_0} \delta^{1/2} \Omega_j^{(n)} + O(\delta). \end{aligned} \tag{60}$$

This map is $O(\delta^{1/2})$ -close to an isometry (a rigid rotation of the variables $(x_j^{(n)}, y_j^{(n)})$). When iterating such maps, a small error added at each iteration will propagate linearly as long as the number of iterations is of order $1/\delta^{1/2}$. Since $K = O(\frac{1}{\delta^{1/2}})$, it follows that the K -th iteration of any map which is $O(\delta^{3/4})$ -close to the time- $\frac{2\pi}{\omega_0}$ map of the flow of (55) is $O(K\delta^{3/4})$ -close to the time- $K\frac{2\pi}{\omega_0}$ map of (55), i.e., $O(\delta^{1/4})$ -close to the K -th iteration of (60).

Thus, to prove the lemma, it is enough to show that the K -th iteration of (60) is $O(\delta^{1/4})$ -close to the time-1 map of (58) (in fact we prove that it is $O(\delta^{1/2})$ -close). We do this by moving to a rotating coordinate frame. Denote

$$\lambda_j = \frac{1}{K} \left\lfloor K \frac{\omega_j}{\omega_0} \right\rfloor, \quad \nu_j = 2\pi K \left(\frac{\omega_j}{\omega_0} - \lambda_j \right)$$

(note that these are the same ν_j as in the statement of the lemma). The rotating coordinate frame corresponds to the new variables:

$$\alpha_j^{(n)}(k) = \phi_j^{(n)}(k) - 2\pi k \lambda_j$$

i.e., at each iteration of the map (60) we subtract $2\pi \lambda_j$ from $\phi_j^{(n)}$. This brings the map (60) to the form:

$$\begin{aligned} \bar{J} &= J - \frac{1}{K} \partial_\psi \hat{U}(\psi) + O(\delta), \\ \bar{\psi} &= \psi + \frac{1}{K} \left(\frac{a}{N} J + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b \hat{I}^{(n)}) R_n \right) + O(\delta), \\ \bar{I}_j^{(n)} &= I_j^{(n)}, \\ \bar{\alpha}_j^{(n)} &= \alpha_j^{(n)} + \frac{1}{K} (\nu_j + \Omega_j^{(n)}) + O(\delta) \end{aligned} \tag{61}$$

(we use here that $\frac{1}{K} = \frac{2\pi}{\omega_0}\delta^{1/2} + O(\delta)$). This is a near-identity map which is $O(\delta)$ -close to the time- $\frac{1}{K}$ map of the system

$$\begin{aligned} \dot{J} &= -\partial_\psi \hat{U}(\psi), \\ \dot{\psi} &= \frac{a}{N}J + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b\hat{I}^{(n)})R_n, \\ \dot{I}_j^{(n)} &= 0, \\ \dot{\alpha}_j^{(n)} &= \nu_j + \Omega_j^{(n)}. \end{aligned} \tag{62}$$

Therefore, the K -th iteration of the map (61) is $O(K\delta)$ - (i.e., $O(\delta^{1/2})$ -) close to the time-1 map of this system. Returning to the non-rotating phases $\phi_j^{(n)}$ does not change the K -th iteration of the map: since $K\lambda_j$ are integers by construction, $\phi_j^{(n)}$ coincides, after the K -th iteration, with $\alpha_j^{(n)}$ modulo 2π , for all j .

It remains to note that system (62) with $\Omega_j^{(n)}$ defined by (56) indeed corresponds to the Hamiltonian (58) (where $\alpha_j^{(n)}$ are the angular variables conjugate to the actions $I_j^{(n)}$). \square

We show next that the twist Assumption SP2 and the KAM assumption IP1 imply:

Lemma 3.4. *At $h = 0$, the system (62) (corresponding to the Hamiltonian (58)) has a positive measure set of KAM tori near its equilibrium at the origin.*

Proof. First, we make a change of coordinates which decouples the (J, ψ) and (\hat{I}, α) degrees of freedom in (58). We achieve this goal by replacing

$$J \rightarrow J - \frac{\sqrt{N}}{a} \sum_{n=1}^N (b\hat{I}^{(n)})R_n, \tag{63}$$

where $R_n = (R_{2,n}, \dots, R_{N,n})$. Then, the right-hand side of the equation for $\dot{\psi}$ will be independent of \hat{I} (i.e., independent of (x, y)). In order to make this a symplectic transformation, we write it as

$$J_j = \tilde{J}_j - \frac{\sqrt{N}}{a} \sum_{n=1}^N \sum_{l=1}^{d-1} b_l I_l^{(n)} R_{j+1,n}, \quad j = 1, \dots, N-1$$

and also transform the α -variables:

$$\alpha_l^{(n)} = \tilde{\alpha}_l^{(n)} + b_l \frac{\sqrt{N}}{a} \sum_{j=1}^{N-1} \psi_j R_{j+1,n}, \quad l = 1, \dots, d-1, \quad n = 1, \dots, N.$$

The simplicity of the transformation $(J, \psi, \hat{I}, \alpha)$ to $(\tilde{J}, \tilde{\psi} = \psi, \tilde{I} = \hat{I}, \tilde{\alpha})$ follows because it is defined by the generating function $\sum_{j=1}^{N-1} ((J_j + \frac{\sqrt{N}}{a} \sum_{n=1}^N \sum_{l=1}^{d-1} b_l I_l^{(n)} R_{j+1,n})\tilde{\psi}_j + \sum_{n=1}^N I_j^{(n)} \tilde{\alpha}_j^{(n)})$.

Performing this change of variables directly in the Hamiltonian (58) at $h = 0$, the new Hamiltonian is (omitting the tilde signs)

$$\begin{aligned}
H = \frac{a}{2N}J^2 + \hat{U}(\psi) - \frac{1}{2a} \left(\sum_{n=1}^N (b\hat{I}^{(n)})R_n \right)^2 + \nu \sum_{n=1}^N \hat{I}^{(n)} + \\
+ \frac{1}{2} \sum_{n=1}^N \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \hat{I}^{(n)} S \hat{I}^{(m)}.
\end{aligned} \tag{64}$$

This is the sum of the Hamiltonian (14) that depends only on J and ψ and describes oscillations around the equilibrium at $(J = 0, \psi = 0)$, and the Hamiltonian

$$H_I = \nu \sum_{n=1}^N \hat{I}^{(n)} - \frac{1}{2a} \left(\sum_{n=1}^N (b\hat{I}^{(n)})R_n \right)^2 + \frac{1}{2} \sum_{n=1}^N \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \hat{I}^{(n)} S \hat{I}^{(m)}, \tag{65}$$

which depends only on \hat{I} variables and describes rotations of the phases α . By the KAM Assumption IP1, the Hamiltonian (14) has a positive measure set of KAM tori near the zero equilibrium. This means we only need to check that the Hamiltonian H_I also has a positive set of KAM tori near the origin, which is proved next, in Lemma 3.5. \square

Lemma 3.5. *The Hamiltonian H_I of (65) satisfies the twist condition.*

Proof. This condition is the requirement that the matrix of second derivatives of H_I with respect to $\hat{I}_j^{(n)}$ is non-degenerate, i.e., the quadratic form

$$-\frac{1}{2a} \left(\sum_{n=1}^N (b\hat{I}^{(n)})R_n \right)^2 + \frac{1}{2} \sum_{n=1}^N \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \hat{I}^{(n)} S \hat{I}^{(m)} \tag{66}$$

is non-degenerate. This is equivalent to the non-degeneracy of the quadratic form

$$\frac{a}{2N}J^2 - \frac{1}{2a} \left(\sum_{n=1}^N (b\hat{I}^{(n)})R_n \right)^2 + \frac{1}{2} \sum_{n=1}^N \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \hat{I}^{(n)} S \hat{I}^{(m)},$$

where we added the dummy variables $J = (J_1, \dots, J_{N-1})^\top$. Replacing

$$J = \tilde{J} + \frac{\sqrt{N}}{a} \sum_{n=1}^N (b\hat{I}^{(n)})R_n$$

(note that this is the inverse of (63)), and omitting the tilde sign, we obtain the quadratic

form

$$\frac{a}{2N} J^2 + \frac{1}{\sqrt{N}} \sum_{n=1}^N (b \hat{I}^{(n)}) (R_n^\top J) + \frac{1}{2} \sum_{n=1}^N \hat{I}^{(n)} \hat{A} \hat{I}^{(n)} - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \hat{I}^{(n)} S \hat{I}^{(m)};$$

Proving its non-degeneracy amounts to showing the non-vanishing of the determinant of the following matrix

$$M = \begin{pmatrix} \begin{pmatrix} \frac{a}{N} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \frac{a}{N} \end{pmatrix} & \frac{1}{\sqrt{N}} R_1 b & \dots & \dots & \dots & \frac{1}{\sqrt{N}} R_N b \\ \frac{1}{\sqrt{N}} b^\top R_1^\top & \hat{A} - S & -S & \dots & \dots & -S \\ \vdots & -S & \hat{A} - S & -S & \dots & -S \\ \vdots & \vdots & \dots & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & \ddots & \vdots \\ \frac{1}{\sqrt{N}} b^\top R_N^\top & -S & \dots & \dots & -S & \hat{A} - S \end{pmatrix}. \quad (67)$$

Let us show that

$$\det M = -\frac{1}{\omega_0^2} \left(\frac{1}{N} \det A \right)^{N-1} \det A_\omega. \quad (68)$$

so, by the single particle twist Assumption, SP3,

$$\det M \neq 0,$$

which will prove the lemma and the theorem.

Recall that R_n 's in the expression (67) are the columns of an orthogonal matrix R without its first row. The first row of R equals to $\frac{1}{\sqrt{N}} (1, \dots, 1)$. Hence, it follows from $R^\top R = id$ that

$$R_n^\top R_m = \begin{cases} -1/N & \text{if } m \neq n, \\ 1 - 1/N & \text{if } m = n, \end{cases} \quad (69)$$

and since $RR^\top = id$

$$R_1 + \dots + R_N = 0. \quad (70)$$

Subtract the last column in formula (67) from each other column, except for the first one.

The resulting matrix

$$\left(\begin{array}{c} \left(\begin{array}{ccc} \frac{a}{N} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \frac{a}{N} \end{array} \right) \\ \frac{1}{\sqrt{N}} b^\top R_1^\top \\ \vdots \\ \vdots \\ \frac{1}{\sqrt{N}} b^\top R_N^\top \end{array} \quad \begin{array}{cccccc} \frac{1}{\sqrt{N}}(R_1 - R_N)b & \frac{1}{\sqrt{N}}(R_2 - R_N)b & \dots & \dots & \frac{1}{\sqrt{N}}R_Nb \\ \hat{A} & 0 & \dots & \dots & -S \\ 0 & \hat{A} & 0 & \dots & -S \\ \vdots & \dots & \ddots & \dots & \vdots \\ 0 & \dots & 0 & \hat{A} & -S \\ -\hat{A} & \dots & \dots & -\hat{A} & \hat{A} - S \end{array} \right)$$

has the same determinant as M . We again get a matrix with the same determinant when, in the last formula, we add all rows, except for the first one, to the last row. The result is

$$\left(\begin{array}{c} \left(\begin{array}{ccc} \frac{a}{N} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \frac{a}{N} \end{array} \right) \\ \frac{1}{\sqrt{N}} b^\top R_1^\top \\ \vdots \\ \vdots \\ \frac{1}{\sqrt{N}} b^\top R_{N-1}^\top \\ 0 \end{array} \quad \begin{array}{cccccc} \frac{1}{\sqrt{N}}(R_1 - R_N)b & \frac{1}{\sqrt{N}}(R_2 - R_N)b & \dots & \dots & \frac{1}{\sqrt{N}}R_Nb \\ \hat{A} & 0 & \dots & \dots & -S \\ 0 & \hat{A} & 0 & \dots & -S \\ \vdots & \dots & \ddots & \dots & \vdots \\ 0 & \dots & 0 & \hat{A} & -S \\ 0 & \dots & \dots & 0 & \hat{A} - NS \end{array} \right).$$

Note that the utmost left bottom block in this matrix equals to $\frac{1}{\sqrt{N}} b^\top (R_1 + \dots + R_N)^\top$ and is zero by (70).

Next, for each $n = 1, \dots, N - 1$ we multiply the first row in this formula by $\frac{\sqrt{N}}{a} b^\top R_n^\top$ and subtract the result from the $(n + 1)$ -th row (so we do not change the first and the last

rows). By (69), this gives us the block-triangular matrix

$$\begin{pmatrix} \left(\begin{array}{ccc} \frac{a}{N} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \frac{a}{N} \end{array} \right) & \frac{1}{\sqrt{N}}(R_1 - R_N)b & \frac{1}{\sqrt{N}}(R_2 - R_N)b & \cdots & \cdots & \frac{1}{\sqrt{N}}R_Nb \\ 0 & \hat{A} - \frac{1}{a}b^\top b & 0 & \cdots & \cdots & -S + \frac{1}{aN}b^\top b \\ \vdots & 0 & \hat{A} - \frac{1}{a}b^\top b & 0 & \cdots & -S + \frac{1}{aN}b^\top b \\ \vdots & \vdots & \cdots & \ddots & \cdots & \vdots \\ \vdots & 0 & \cdots & 0 & \hat{A} - \frac{1}{a}b^\top b & -S + \frac{1}{aN}b^\top b \\ 0 & 0 & \cdots & \cdots & 0 & \hat{A} - NS \end{pmatrix}.$$

By construction, its determinant equals to the determinant of M , which gives

$$\det M = \left(\frac{a}{N}\right)^{N-1} \det(\hat{A} - \frac{1}{a}b^\top b)^{N-1} \det(\hat{A} - NS).$$

Now, formula (68) follows, since

$$\det A = \det \begin{pmatrix} a & b \\ b^\top & \hat{A} \end{pmatrix} = \det \begin{pmatrix} a & 0 \\ b^\top & \hat{A} - \frac{1}{a}b^\top b \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{a}b \\ 0 & I \end{pmatrix} = a \det(\hat{A} - \frac{1}{a}b^\top b),$$

and, by (59),

$$\begin{aligned} -\omega_0^2 \det(\hat{A} - NS) &= -\omega_0^2 \det\left(\hat{A} - \frac{1}{\omega_0}(\hat{\omega}^\top b + b^\top \hat{\omega} - \frac{a}{\omega_0} \hat{\omega}^\top \hat{\omega})\right) = \det \begin{pmatrix} -\frac{\omega_0^2}{a} & \omega_0 & \hat{\omega} \\ 0 & a & b \\ 0 & 0 & \hat{A} + \frac{a}{\omega_0^2}(\hat{\omega}^\top - \frac{\omega_0}{a}b^\top)\hat{\omega} - \frac{\hat{\omega}^\top b}{\omega_0} \end{pmatrix} \\ &= \det \begin{pmatrix} -\frac{\omega_0^2}{a} & \omega_0 & \hat{\omega} \\ 0 & a & b \\ 0 & b^\top + \frac{a}{\omega_0}(\hat{\omega}^\top - \frac{\omega_0}{a}b^\top) & \hat{A} + \frac{a}{\omega_0^2}(\hat{\omega}^\top - \frac{\omega_0}{a}b^\top)\hat{\omega} \end{pmatrix} = \det \begin{pmatrix} -\frac{\omega_0^2}{a} & \omega_0 & \hat{\omega} \\ 0 & a & b \\ \hat{\omega}^\top - \frac{\omega_0}{a}b^\top & b^\top & \hat{A} \end{pmatrix} = \\ &= \det \begin{pmatrix} 0 & \omega_0 & \hat{\omega} \\ \omega_0 & a & b \\ \hat{\omega}^\top & b^\top & \hat{A} \end{pmatrix} = \det A_\omega. \end{aligned}$$

□

This completes the proof of Lemma 3.4, showing that the Hamiltonian system (58) has a positive measure set of KAM tori.

Now, the claim of Theorem 1 follows. Indeed, as KAM-tori persist at small perturbations,

Lemmas 3.2 and 3.3 imply that system (54) also has a positive measure set of KAM-tori on every energy level $H = h$ with small h . Since the system defined by (54) is smoothly conjugate to the original N -particle system (42) near $\mathbf{L}^*(\theta_{min})$, with a scaling factor $O(\delta^{1/4})$, the KAM theory implies that there are quasi-periodic orbits that are $O(\delta^{1/4})$ -close to $\mathbf{L}^*(\theta_{min})$. By (57), on such tori, the return time to the cross-section $\varphi = 0$ is $O(\delta^{1/2})$ -close to $\frac{2\pi}{\omega_0}$. It follows that the averaged return time, $\frac{2\pi}{\bar{\omega}}$ is also $O(\delta^{1/2})$ -close to $\frac{2\pi}{\omega_0}$, completing the proof of Theorem 1.

4 Mutually repelling particles in a container

The motion of mutually repelling particles confined in a bounded domain D is described by the Hamiltonian (19). In the limit of high energy per particle (i.e. for $\delta = \frac{1}{2h} \rightarrow 0$), one can view the system as a set of weakly interacting particles in a steep billiard-like potential as described by (20), with the single-particle dynamics governed by (21).

The singularity of the single-particle system at $\delta = 0$ makes the proof of Theorem 3 more involved than for Theorem 1. The outline of the proof is as follows. In Section 4.1, we study the single-particle system (21) in the small δ limit. Recall that this system has an elliptic periodic orbit for $\delta = 0$, and hence, by assumptions Box1 and Box2, this orbit persists also for sufficiently small δ . By studying the singular behavior near impacts, we construct a transformation to action-angle coordinates near this periodic orbit, with a singularity of the transformation near the billiard boundary. We prove that the Hamiltonian expressed in these coordinates has a smooth limit at $\delta = 0$ (see Section 4.1.3). We then show that for all sufficiently small $\delta \geq 0$ this periodic orbit satisfies Assumptions SP1-SP3 (Lemma 4.4). In Section 4.2, we study the multi-particle dynamics. Here, using the analysis of Section 4.1, we show that the return map to a cross-section at which all particles are away from the billiards' boundary is not singular at $\delta = 0$ and is $o(\delta^{1/2})$ -close to the return map of a certain truncated system (system (100), see Lemma 4.5). We then analyze the return map of the truncated system by averaging (Lemma 4.7). Due to the singular nature of the impacts, we use the pair-wise structure of the interaction terms and not the Fourier expansion which was used in Lemma 3.1. We then establish that the return map of the truncated system (100) is close to that of the truncated averaged system (117). This system is of the same form as the truncated averaged system considered in Section 3 (cf. (46)), with the only difference that the coefficients now depend on δ . Since the dependence of the coefficients is non-singular (continuous) for all $\delta \geq 0$, we can conclude the proof as in Theorem 1.

4.1 A single particle at high energy.

Here we study the solutions of (21) for small δ . Away from the boundary, $\delta V(q)$ is uniformly small, so the trajectories follow closely the corresponding billiard trajectories. Near the billiard boundary, a more precise analysis is needed.

4.1.1 The boundary layer dynamics

By the boundary layer, we mean a sufficiently small neighborhood of the boundary of D where we have $V(q) = 1/Q(q)^\alpha$, with $Q(q)$ measuring the distance to the boundary of D (see Assumption BD1).

Lemma 4.1. *Take a small neighborhood of a regular point $M \in \partial D$. Then, one can define functions \tilde{q} , \tilde{p} , q_{impact} , t_{in} such that the following holds. Given any initial condition (q_0, p_0) which is close to a regular impact (i.e., q_0 is in the small neighborhood of M but is bounded away from ∂D , and $p_0 \cdot \nabla Q(M)$ is negative and bounded away from zero), the trajectory $(q(t, q_0, p_0), p(t, q_0, p_0))$ of (q_0, p_0) can be written, in the boundary layer, in the following form:*

$$\begin{aligned} q(t, q_0, p_0) &= q_{\text{impact}} + \delta^{1/\alpha} \tilde{q}(t_s, q_0, p_0; \delta), \\ p(t, q_0, p_0) &= \tilde{p}(t_s, q_0, p_0; \delta), \end{aligned} \tag{71}$$

where t_s denotes the rescaled time

$$t_s = \frac{t - t_{\text{in}}(q_0, p_0; \delta)}{\delta^{1/\alpha}}. \tag{72}$$

The functions \tilde{p}, \tilde{q} depend smoothly on (t_s, q_0, p_0) and, along with the derivatives, depend continuously on δ for all $\delta \geq 0$: the function \tilde{p} , along with the derivatives, is uniformly bounded and uniformly continuous for all t_s and $\delta \geq 0$, and the function \tilde{q} is given by

$$\tilde{q}(t_s, q_0, p_0; \delta) = \tilde{q}(0, q_0, p_0; \delta) + \int_0^{t_s} \tilde{p}(u, q_0, p_0; \delta) du, \tag{73}$$

where $\tilde{q}(0, q_0, p_0; \delta)$ depends continuously on δ , along with the derivatives, for all $\delta \geq 0$. The function $q_{\text{impact}}(q_0, p_0)$ is smooth and independent of δ , and is determined by the billiard impact event:

$$q_{\text{impact}}(q_0, p_0) \in \partial D. \tag{74}$$

The function $t_{\text{in}}(q_0, p_0; \delta)$ depends smoothly on (q_0, p_0) and, along with the derivatives, depends continuously on δ for all $\delta \geq 0$. In the limit $\delta \rightarrow 0$, the trajectory approaches the billiard trajectory, i.e.,

$$q_0 + t_{\text{in}} p_0 = q_{\text{impact}} \text{ for } \delta = 0, \tag{75}$$

and

$$\lim_{t_s \rightarrow -\infty} \tilde{p}(t_s, q_0, p_0; 0) = p_0, \quad \lim_{t_s \rightarrow +\infty} \tilde{p}(t_s, q_0, p_0; 0) = p_0 - 2(p_0 \cdot \vec{n})\vec{n} \quad (76)$$

(the billiard reflection law) where \vec{n} is the outer normal to ∂D at the impact point $q_{\text{impact}}(q_0, p_0; 0)$.

Proof. We follow the strategy of [37, 51]. Put the origin of the coordinate system at the point M , and let q_{\parallel} denote coordinates corresponding to directions tangent to the boundary at the origin, and the q_{\perp} -coordinate axis be orthogonal to the boundary at M . So, near the origin, we can write

$$Q(q_{\parallel}, q_{\perp}) = q_{\perp} + O(q_{\parallel}^2 + q_{\perp}^2). \quad (77)$$

The equation of motion for the single-particle Hamiltonian (21) are

$$\begin{aligned} \frac{d}{dt} q_{\parallel} &= p_{\parallel}, & \frac{d}{dt} q_{\perp} &= p_{\perp}, \\ \frac{d}{dt} p_{\parallel} &= \frac{\alpha \delta \nabla_{\parallel} Q}{Q^{\alpha+1}}, & \frac{d}{dt} p_{\perp} &= \frac{\alpha \delta \nabla_{\perp} Q}{Q^{\alpha+1}}. \end{aligned} \quad (78)$$

We take a small $\eta > 0$ and consider an $O(\eta)$ -neighborhood of M . In this region, $q = O(\eta)$, and $\nabla Q(q) = (\nabla_{\parallel} Q, \nabla_{\perp} Q)_q = (O(\eta), 1 + O(\eta))$. Then, we see that the value of p_{\perp} monotonically increases with time and the change in p_{\parallel} is much smaller than the change of p_{\perp} (because $dp_{\parallel}/dp_{\perp} = O(\eta)$ which is small). Since the range of possible values of p_{\perp} is bounded by the energy conservation, it follows that p_{\parallel} is an almost conserved quantity (can change at most by $O(\eta)$).

Take a sufficiently large constant K . Define the outer boundary layer as the region $\delta^{1/\alpha} K < Q(q) < \eta$ (in particular, we consider sufficiently small δ so that $\delta^{1/\alpha} \ll \eta$). For sufficiently small δ , the initial value q_0 belongs to this region. In this outer layer, the value of the potential energy $\frac{\delta}{Q^{\alpha}}$ is small of order $O(K^{-\alpha})$, so the maximal possible change in the kinetic energy is $O(K^{-\alpha})$ as long as the trajectory stays in the layer. Thus, the velocity vector $p = \dot{q}$ remains almost constant, $p(t) = p_0 + O(K^{-\alpha} + \eta)$, by the approximate conservation of the kinetic energy and of p_{\parallel} . It follows that the trajectory is close to a straight line (i.e., to the billiard trajectory) and $q(t)$ moves inward, towards the impact. Therefore, there exists some time, $t_{in}(q_0, p_0; \delta) = O(\eta)$ at which the trajectory crosses the surface $S_{K,\delta} : Q(q) = \delta^{1/\alpha} K$.

Let us obtain more precise estimates for $(q(t), p(t))$ for $t \leq t_{in}$. On this time interval, $\frac{dQ}{dt} = \nabla Q(q) \cdot p = p_{\perp}(t) + O(\eta) = p_{\perp}(0) + O(K^{-\alpha} + \eta) < 0$. Hence, we can choose Q as a new time. In fact, it is more convenient to choose $s = \delta^{-1/\alpha} Q$ as the new rescaled time, so the equations become:

$$\begin{aligned} \frac{dq}{ds} &= \delta^{1/\alpha} \frac{p}{f(q,p)}, \\ \frac{dp}{ds} &= \frac{\alpha}{s^{\alpha+1}} \frac{\nabla Q(q)}{f(q,p)}, \\ \frac{dt}{ds} &= \delta^{1/\alpha} \frac{1}{f(q,p)}, \end{aligned} \quad (79)$$

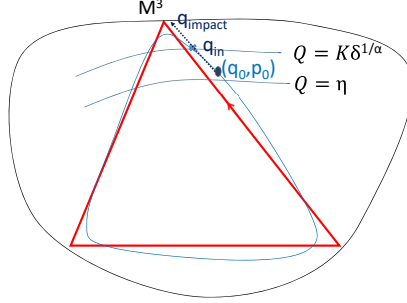


Figure 6: Boundary layers near impact for small δ . The outer boundary layer lies between the level sets $Q = \eta$ and $Q = K\delta^{1/\alpha}$ whereas the inner layer is defined by $Q < K\delta^{1/\alpha}$.

where $f = \nabla Q(q) \cdot p$ is a smooth function of (q, p) . Hence,

$$\begin{aligned}
 q(s) &= q_0 + \delta^{1/\alpha} \int_{s_0}^s \frac{p(\sigma)}{f(q(\sigma), p(\sigma))} d\sigma = q_0 + \int_{\delta^{1/\alpha}s_0}^{\delta^{1/\alpha}s} \frac{p(\sigma)}{f(q(\sigma), p(\sigma))} d(\delta^{1/\alpha}\sigma), \\
 p(s) &= p_0 + \int_{s_0}^s \frac{\nabla Q(q(\sigma), p(\sigma))}{f(q(\sigma), p(\sigma))} \frac{\alpha d\sigma}{\sigma^{\alpha+1}} = p_0 - \int_{s_0^{-\alpha}}^{s^{-\alpha}} \frac{\nabla Q(q(\sigma), p(\sigma))}{f(q(\sigma), p(\sigma))} d(\sigma^{-\alpha}), \\
 t(s) &= \delta^{1/\alpha} \int_{s_0}^s \frac{1}{f(q(\sigma), p(\sigma))} d\sigma = \int_{\delta^{1/\alpha}s_0}^{\delta^{1/\alpha}s} \frac{d(\delta^{1/\alpha}\sigma)}{f(q(\sigma), p(\sigma))}
 \end{aligned} \tag{80}$$

where $s \in [K, \delta^{-1/\alpha}\eta]$. The solution of this system of integral equations is obtained by the contraction mapping principle: the integrands are bounded with all derivatives with respect to q, p and the integration intervals are small (of order $O(\eta)$ in the first and third equations and $O(K^{-\alpha})$ in the second equation). It follows that, for any fixed s_0, s in the outer boundary layer, we have smooth dependence of (q, p, t) on (q_0, p_0) , for all small δ , including the limit $\delta = 0$. The smoothness with respect to s_0, s follows from the system (79), as s is bounded away from zero.

We denote the solution of this system by $(q^{int}(s, q_0, p_0; \delta), p^{int}(s, q_0, p_0; \delta), t^{int}(s, q_0, p_0; \delta))$. Note that $(q^{int}(s_0, q_0, p_0; \delta), p^{int}(s_0, q_0, p_0; \delta), t^{int}(s_0, q_0, p_0, t_0; \delta)) = (q_0, p_0, 0)$. Since the time t_{in} corresponds to the time instance the trajectory hits the cross-section $S_{K,\delta}$, i.e., it corresponds to $s = \delta^{-1/\alpha}Q = K$, we obtain that

$$t_{in}(q_0, p_0; \delta) = t^{int}(K, q_0, p_0; \delta).$$

We also define

$$q_{in}(q_0, p_0; \delta) = q^{int}(K, q_0, p_0; \delta),$$

i.e., the point where the orbit of (q_0, p_0) hits $S_{K,\delta} : Q(q) = K\delta^{1/\alpha}$. As we have shown, these are smooth functions of (q_0, p_0) , uniformly for all $\delta \geq 0$.

Introducing the rescaled time t_s by formula (72), we find that $\frac{\partial}{\partial s} t_s = \frac{\partial}{\partial s} \frac{t^{int}}{\delta^{1/\alpha}} = \frac{1}{f(q^{int}, p^{int})}$, which is bounded with all derivatives and is bounded away from zero. Hence, $s(t_s, q_0, p_0; \delta)$ is a smooth function of its arguments for all $\delta \geq 0$ (large s corresponds to approaching the outer boundary of the boundary layer, where $t_s \rightarrow -\infty$). Therefore, in the outer boundary layer, the function

$$\tilde{p}(t_s, q_0, p_0; \delta) = p^{int}(s(t_s, q_0, p_0; \delta), q_0, p_0; \delta) \quad (81)$$

is a smooth function of its arguments. Since $\dot{q} = p$, and $q(t_{in}, q_0, p_0) = q_{in}$, we have

$$q(t; q_0, p_0) = q_{in} + \int_{t_{in}}^t \tilde{p}(t_s, q_0, p_0; \delta) dt = q_{impact} + (q_{in} - q_{impact}) + \delta^{1/\alpha} \int_0^{t_s} \tilde{p}(u, q_0, p_0; \delta) du.$$

Note that q_{in} is the point on $S_{K,\delta}$, i.e., $Q(q_{in}) = K\delta^{1/\alpha}$. We also have $Q(q_{impact}) = 0$ (see (74)), hence $Q(q_{in}) - Q(q_{impact}) = K\delta^{1/\alpha}$. Since the gradient of Q is bounded away from zero, it follows that $(q_{in} - q_{impact})$ is $\delta^{1/\alpha}$ times a smooth function of (q_0, p_0) , continuously depending on $\delta \geq 0$ with all derivatives. Hence it can be incorporated into \tilde{q} :

$$\tilde{q}(t_s, q_0, p_0; \delta) = \delta^{-1/\alpha}(q_{in} - q_{impact}) + \int_0^{t_s} \tilde{p}(u, q_0, p_0; \delta) du.$$

As we see, the claim of the lemma, including formulas (71),(73), follows for the initial segment of the orbit (i.e., as long as it stays in the outer layer).

Let us now prove formula (75). By (71), (72), (73), we have

$$q_0 + p_0 t_{in}(q_0, p_0; \delta) = q_{impact} + O(\delta^{1/\alpha}) + \delta^{1/\alpha} \int_0^{-\delta^{-1/\alpha} t_{in}} (\tilde{p}(t_s, q_0, p_0; \delta) - p_0) dt_s.$$

This implies (75) because the last term in the above formula tends to zero, along with all derivatives, as $\delta \rightarrow 0$. Indeed, by (81) and by the second equation of (80) we have

$$\tilde{p}(t_s, q_0, p_0; \delta) - p_0 = O(s^{-\alpha}), \quad (82)$$

along with derivatives up to any given order. So, since $\frac{dt_s}{ds} = \frac{dt_s}{dt} \frac{dt}{ds}$ is uniformly bounded with derivatives (by (72) and the third equation of (79)), we have that (recall that s_0 is bounded

by $\eta\delta^{-1/\alpha}$):

$$\delta^{1/\alpha} \int_K^{s_0} (\tilde{p}(t_s, q_0, p_0; \delta) - p_0) \frac{dt_s}{ds} ds = \begin{cases} \delta^{1/\alpha} O(s_0^{1-\alpha}) = O(\delta), & \alpha < 1, \\ \delta^{1/\alpha} O(\ln(s_0)) = O(\delta^{1/\alpha} \ln \delta), & \alpha = 1, \\ O(\delta^{1/\alpha}), & \alpha > 1, \end{cases} \quad (83)$$

as required.

Next, we study the trajectory in the inner layer $Q(q) \leq \delta^{1/\alpha} K$. Here, p_\perp changes rapidly, so the trajectory quickly exits this inner layer, intersects the surface $S_{K,\delta}$ again, and returns back to the outer layer. After the rescaling $\bar{q} = \delta^{-1/\alpha}(q - q_{in}), t_s = \delta^{-1/\alpha}(t - t_{in})$, the Hamiltonian (21) becomes

$$H_0 = \frac{p^2}{2} + \bar{Q}(\bar{q}, \delta)^{-\alpha} \quad (84)$$

where

$$\bar{Q}(\bar{q}, \delta) = \frac{Q(q_{in} + \delta^{1/\alpha} \bar{q})}{\delta^{1/\alpha}}. \quad (85)$$

Note that $\bar{Q}(\bar{q}, \delta)$ is a smooth function of \bar{q} with bounded derivatives for all $\delta \geq 0$. By construction, the inner layer is given by $K \geq \bar{Q} \geq H_0^{-1/\alpha} > 0$ (where H_0 is the conserved energy, see (84)).

The rescaled system, as given by the Hamiltonian (84), is

$$\frac{d\bar{q}}{dt_s} = p, \quad \frac{dp_\parallel}{dt_s} = \frac{\alpha \nabla_\parallel \bar{Q}}{\bar{Q}^{\alpha+1}}, \quad \frac{dp_\perp}{dt_s} = \frac{\alpha \nabla_\perp \bar{Q}}{\bar{Q}^{\alpha+1}}, \quad (86)$$

where $\nabla_\parallel \bar{Q} = \nabla_\parallel Q = O(\|q_{in}\| + \delta^{1/\alpha})$ and $\nabla_\perp \bar{Q} = \nabla_\perp Q = 1 + O(K\delta^{1/\alpha}) > 0$. Since \bar{Q} is uniformly bounded in the inner layer, it follows that $\frac{d}{dt_s} p_\perp$ is positive and bounded away from zero. By the conservation of energy, p_\perp cannot grow unbounded, hence the orbit must leave the inner layer in a finite time, which we denote $t_{s,out}$. This time is bounded for all small $\delta \geq 0$ (so the unscaled passage time is of order $O(\delta^{1/\alpha})$).

The system (86) is well-defined at $\delta = 0$, so the solution on any finite interval of the integration time t_s is a smooth function of the initial conditions and parameters, continuously depending on δ for all small $\delta \geq 0$. Note that the initial condition at $t_s = 0$ is $\bar{q} = 0$, $p = p^{int}(K, q_0, p_0; \delta)$; the right-hand side also depends smoothly on the value of q_{in} , which is a smooth function of q_0, p_0 . Thus, we have a smooth dependence on q_0, p_0 and the scaled time t_s for all small $\delta \geq 0$, i.e., the claim of the lemma continues to hold as long as the solution is in the inner layer.

Let us show that the exit time $t_{s,out}$ is a smooth function of the initial conditions and parameters of the system, i.e., it is a smooth function of q_0 and p_0 . This moment of time

corresponds to arriving at the cross-section $S_{K,\delta} : \bar{Q} = K$, so we just need to show that $\frac{d}{dt_s} \bar{Q}$ is bounded away from zero. To do that, note that since the time in the boundary layer is bounded, the change in p_{\parallel} is small, of order $O(\nabla_{\parallel} \bar{Q}) = O(\|q_{in}\| + \delta^{1/\alpha})$. Since the energy (84) is conserved and the value of \bar{Q} at the entrance and the exit from the boundary layer is the same, the kinetic energy $\frac{1}{2}(p_{\parallel}^2 + p_{\perp}^2)$ is also the same. Hence $|p_{\perp}|$ at the moment of exit is $O(\|q_{in}\| + \delta^{1/\alpha})$ -close to the value of $|p_{\perp}|$ at the moment of entrance, so it is $O(K^{-\alpha})$ -close to $|p_{0\perp}|$ by (82) (the sign of p_{\perp} must change since the orbit is going away from the billiard boundary now). It follows that

$$\frac{d}{dt_s} \bar{Q} = \nabla_{\perp} Q \cdot p_{\perp} + \nabla_{\parallel} Q \cdot p_{\parallel} = -p_{0\perp} + O(\|q_{in}\| + \delta^{1/\alpha} + K^{-\alpha}) > 0$$

is bounded away from zero, as required.

As $t_{s,out}$ depends smoothly on q_0 and p_0 for all $\delta \geq 0$, the values of $q = q_{out}$ and $p = p_{out}$ at the moment of exiting the inner layer also depend smoothly on q_0 and p_0 . Note that we have just shown that

$$p_{out} = (p_{0\parallel}, -p_{0\perp}) + O(\|q_{in}\| + \delta^{1/\alpha} + K^{-\alpha}). \quad (87)$$

Once the trajectory crosses $S_{K,\delta}$ towards the outer boundary layer (i.e., $p_{\perp} > 0$ now), we can again use the integral equations (80) to establish the smooth dependence on the initial conditions (q_{out}, p_{out}) - hence on (q_0, p_0) - and the scaled time. So, the solutions in this final segment also satisfy the claim of the lemma.

It remains to establish the reflection law (76). For a given initial condition, we choose the origin of coordinates to be the billiard impact point. The billiard reflection law then is that p_{\parallel} remains the same and p_{\perp} changes sign. By (82), we have that

$$\tilde{p} = (p_{0\parallel}, p_{0\perp}) + O(K^{-\alpha})$$

before entering the inner layer and, taking into account the change in p in the inner layer, as given by (87) we find that in the limit $\delta \rightarrow 0$

$$\tilde{p} = (p_{0\parallel}, -p_{0\perp}) + O(K^{-\alpha})$$

after exiting the inner layer (we have $q_{in} \rightarrow 0$ as $\delta \rightarrow 0$ because we put the coordinate origin at the billiard impact point $q_{impact} = \lim_{\delta \rightarrow 0} q_{in}$). Thus, except for the bounded interval of the rescaled time t_s for which the orbit is in the inner layer, the deviation from the billiard reflection law is bounded by $O(K^{-\alpha})$ as $\delta \rightarrow 0$. Since K can be chosen as large as we want,

and the result cannot depend on K , this means that the deviation from the billiard law in the limit $\delta = 0$ is zero, i.e., the billiard reflection law is approached indeed. \square

4.1.2 Flow-box coordinates in the boundary layer

Consider a billiard trajectory near a regular impact point $M \in \partial D$.

Lemma 4.2. *One can choose cross-sections S_δ^- and S_δ^+ in the phase space such that the billiard trajectory intersects, transversely, the cross-section S_0^- before the impact (at a point (q'_0, p'_0)) and the cross-section S_0^+ after the impact, and the following holds. For all $(q_0, p_0) \in S_\delta^-$ the impact time $t_{in}(q_0, p_0; \delta)$ of Lemma 4.1 is constant and equal to $t_{in}(q'_0, p'_0; \delta) > 0$, and the flight time from S_δ^- to S_δ^+ is constant for the orbits of system (21) and equals to $2t_{in}(q'_0, p'_0; \delta)$. Such cross-sections are bounded away from the billiard boundary and depend continuously on δ . The Poincaré map $S_\delta^- \rightarrow S_\delta^+$ by the orbits of the system tends, with all derivatives, to the Poincaré map $S_0^- \rightarrow S_0^+$ by the billiard flow.*

Proof. For the billiard flow, the function $t_{in}(q_0, p_0; 0)$ satisfies $Q(q_0 + p_0 \cdot t_{in}) = 0$. Since the impact at point M is regular, we have $p_0 \cdot \nabla Q(q_0) \neq 0$, hence the equation $t_{in}(q, p; 0) = t_{in}(q'_0, p'_0; 0)$ defines a smooth hypersurface, which is the cross-section S_0^- . By the continuous dependence of $t_{in}(q, p; \delta)$ and its derivatives on δ , the equation $t_{in}(q, p; \delta) = t_{in}(q'_0, p'_0; \delta)$ defines a smooth hypersurface, which is the cross-section S_δ^- .

To satisfy the lemma, we must choose S_δ^+ as the image of S_δ^- by the time- $2t_{in}$ flow map. Therefore, by Lemma 4.1, the cross-section S_δ^+ consists of the points (q, p) satisfying

$$q = q_{impact} + \delta^{1/\alpha} \tilde{q}(t_{in} \delta^{-1/\alpha}, q_0, p_0; \delta), \quad p = \tilde{p}(t_{in} \delta^{-1/\alpha}, q_0, p_0; \delta), \quad (88)$$

where (q_0, p_0) is in S_δ^- . Formula (88) defines the Poincaré map $S_\delta^- \rightarrow S_\delta^+$. By (73),

$$q = q_{impact} + p t_{in} + O(\delta^{1/\alpha}) + \delta^{1/\alpha} \int_0^{t_{in} \delta^{-1/\alpha}} (\tilde{p}(u, q_0, p_0; \delta) - p) du.$$

As in Lemma 4.1, it follows (similar to (83)) that the last term in this formula tends to zero as $\delta \rightarrow 0$, along with all derivatives with respect to (q_0, p_0) . Therefore, q and p in (88) have a well-defined limit as $\delta \rightarrow 0$, and S_δ^+ tends to S_0^+ , the time- $2t_{in}$ image of S_0^- by the billiard flow. For regular impacts, S_0^+ is a well-defined smooth hypersurface and is bounded away from the boundary. By continuity, the same is true for S_δ^+ for all small $\delta \geq 0$. \square

Next, we introduce flow-box coordinates [1] for the union of trajectory segments that are close to the billiard trajectory (q^*, p^*) in the boundary layer. Precisely, let \mathcal{U}_δ denote the union of the segments of trajectories of system (21) that cross S_δ^- at $t = 0$ and correspond

to an open time-interval containing the flight-time interval $[0, 2t_{in}]$ (at $\delta = 0$, we take \mathcal{U}_0 as the union of the corresponding billiard trajectories).

Lemma 4.3. *In \mathcal{U}_δ , one can make a symplectic change of coordinates $(q, p) \rightarrow (\tau, \mathcal{E}, \mathcal{P}) \in \mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{R}^{d-2}$, such that (τ, \mathcal{E}) is the symplectic time-energy pair, $\tau = 0$ at S_δ^- and $\tau(q, p; \delta) = 2t_{in}$ at S_δ^+ , and the system in \mathcal{U}_δ acquires the form*

$$\dot{\tau} = 1, \quad \dot{\mathcal{E}} = 0, \quad \dot{\mathcal{P}} = 0. \quad (89)$$

The coordinate transformation is smooth and depends continuously on δ , along with the derivatives, for $\delta > 0$. Moreover, uniformly for all $\delta \geq 0$, the variable q depends C^∞ -smoothly on $(\mathcal{E}, \mathcal{P})$ for every fixed τ , and it depends continuously on τ and δ , along with all derivatives with respect to $(\mathcal{E}, \mathcal{P})$.

Away from the billiard boundary (in particular, near S_δ^- and S_δ^+), the transition maps between the coordinates (q, p) and $(\tau, \mathcal{E}, \mathcal{P})$ are smooth, and depend on δ continuously, in C^∞ , for all $\delta \geq 0$.

Proof. We use the standard flow-box construction: Given a point in \mathcal{U}_δ , we take the value of the Hamiltonian (21) at this point and define \mathcal{E} as this value minus $\frac{1}{2}$; we define τ as the time it takes a trajectory of system (21) to reach this point from S_δ^- . The coordinates \mathcal{P} are taken constant along the trajectory and, hence, equal to the coordinates \mathcal{P} of the intersection point of the trajectory with S_δ^- . One can choose coordinates \mathcal{P} on S_δ^- such that the resulting coordinate system $(\tau, \mathcal{E}, \mathcal{P})$ is symplectic in \mathcal{U}_δ [1]. Formula (89) is immediate from the construction (recall that the energy \mathcal{E} is conserved).

Let us examine the regularity of the coordinate transformation in the limit $\delta \rightarrow 0$. Away from the billiard boundary, the flow of (21) has a regular limit (the billiard flow). It follows that if we remove from \mathcal{U}_δ the points whose q -component belongs to a given small neighborhood of the boundary, then the resulting set will have two connected components for all sufficiently small δ . One component is comprised by trajectory pieces that intersect S_δ^- , and the other by pieces that intersect S_δ^+ . It is immediate by construction that in the first connected component the flow-box coordinate transformation is regular up to $\delta = 0$, as required.

For each point of the second component, one has well-defined (for all $\delta \geq 0$) correspondence between the (q, p) coordinates of the point and the flow-box coordinates $(\tau', \mathcal{E}', \mathcal{P}')$, where $(\mathcal{E}', \mathcal{P}')$ are the coordinates of the intersection of the orbit of the point with S_δ^+ and τ' is the time the orbit needs to arrive to S_δ^+ . We have $\tau = 2t_{in} - \tau'$, so since t_{in} is a constant, it follows that to establish the regularity of the transformation $(q, p) \rightarrow (\tau, \mathcal{E}, \mathcal{P})$ in

the second connected component, we just need to show the regularity of the transformation $(\mathcal{E}', \mathcal{P}') \rightarrow (\mathcal{E}, \mathcal{P})$ for the points on the cross-section S_δ^+ only.

It remains to note that the coordinates $(\mathcal{E}, \mathcal{P})$ stay constant along a trajectory, so for any point on S_δ^+ the correspondence $(\mathcal{E}, \mathcal{P}) \rightarrow (\mathcal{E}', \mathcal{P}')$ defines the Poincaré map $S_\delta^- \rightarrow S_\delta^+$ whose regularity for all $\delta \geq 0$, is established by Lemma 4.2.

We have shown the regularity of the flow-box coordinates away from the billiard boundary. Now, to finish the lemma, we discuss the dependence of q on the flow-box coordinates. Since τ is the time variable and $(\mathcal{E}, \mathcal{P})$ give the initial conditions on the cross-section $S_{\delta,j}^-$, we can write (71) as

$$q = q_{\text{impact}} + \delta^{1/\alpha} \tilde{q}\left(\frac{\tau - t_{in}}{\delta^{1/\alpha}}, \mathcal{E}, \mathcal{P}; \delta\right). \quad (90)$$

Since the value of t_{in} is independent of $(\mathcal{E}, \mathcal{P}) \in S_{\delta,j}^-$, the required regularity of q follows immediately from Lemma 4.1. \square

4.1.3 The periodic orbit L_δ^*

Choose, near each regular impact point M^j , $j = 1, \dots, k^*$, of the billiard orbit L^* , the cross-sections $S_{\delta,j}^-$ and $S_{\delta,j}^+$ as in Lemma 4.2. The cross-sections depend continuously on δ and stay bounded away from the billiard boundary. Therefore, in a neighborhood of the piece of L^* between the consecutive cross-sections $S_{\delta,j}^+$ and $S_{\delta,j+1}^-$, the potential tends to zero as $\delta \rightarrow 0$ and the motion tends to the constant speed motion, as in the billiard. Hence, for the orbits of system (21) the flight time from $S_{\delta,j}^+$ to $S_{\delta,j+1}^-$ tends to the billiard flight time and the Poincaré map $S_{\delta,j}^+ \rightarrow S_{\delta,j+1}^-$ tends to the corresponding Poincaré map of the billiard flow, along with the derivatives with respect to the initial conditions.

By Lemma 4.2, the same is true for the Poincaré map and the flight time between $S_{\delta,j}^-$ and $S_{\delta,j}^+$ (the flight time does not depend on initial conditions and equals to $2t_{in}$, so it tends to the billiard flight time by Lemma 4.1). Thus, we obtain that the return map to $S_{\delta,1}^-$ for the flow (21) is C^r close, for all r , to the return map of the billiard flow (the return map is the composition of the Poincaré maps $S_{\delta,j}^- \rightarrow S_{\delta,j}^+$ and $S_{\delta,j}^+ \rightarrow S_{\delta,j+1}^- \text{ mod } k^*$ for $j = 1, \dots, k^*$). Also, the return time to $S_{\delta,1}^-$ is C^r close to the return time for the billiard flow.

By Assumption BD2, the intersection point of L^* with $S_{0,1}^-$ is a non-degenerate non-resonant (up to order 4) elliptic fixed point of the billiard return map at the energy level $H = \frac{1}{2}$ (this corresponds to the motion with the speed 1). Such fixed points persist at small perturbations, so the return map to $S_{\delta,1}^-$ also has a KAM-non-degenerate elliptic fixed point at the energy level to $H = \frac{1}{2}$ for all small δ .

This gives us an elliptic periodic orbit L_δ^* of system (21) such that $L_\delta^* \cap S_{\delta,1}^-$ tends to $L^* \cap S_{0,1}^-$ as $\delta \rightarrow 0$. By the continuous dependence of the return time on δ , the period of L_δ^* tends to the

period of the billiard orbit L^* . Recall that we use the notation $L^* = (q^*(\omega_0 t), p^*(\omega_0 t))$, where $\omega_0 = \frac{2\pi}{|L^*|}$ and (q^*, p^*) are 2π -periodic functions, q^* is continuous and piece-wise linear, and p^* is discontinuous and piece-wise constant. We denote $L_\delta^* = (q_\delta^*(\omega_0(\delta)t), p_\delta^*(\omega_0(\delta)t))$ for some smooth 2π -periodic functions (q_δ^*, p_δ^*) . Here, $\omega_0(\delta) \rightarrow \omega_0$ as $\delta \rightarrow 0$, and $(q_\delta^*, p_\delta^*) \rightarrow (q^*, p^*)$, with derivatives, if q^* stays away from the billiard boundary. Since q_δ^* is uniformly Lipschitz and the time spent in the boundary layer is small, it follows that $q_\delta^* \rightarrow q^*$ in C^0 for all t .

The same holds true in every energy level close to $H = \frac{1}{2}$. The billiard has a periodic orbit $L^*(E)$ in the energy level $H = E$ which follows the same path as L^* in the q -space, with the speed $\|p\| = \sqrt{2E}$ and period $T(E) = \frac{|L^*|}{\sqrt{2E}}$. By the same arguments as above, the system (21) for all sufficiently small δ has an elliptic periodic orbit $L_\delta^*(E)$ in the energy level $H = E$, and the family of the orbits $L_\delta^*(E)$ approaches $L^*(E)$ as $\delta \rightarrow 0$. In particular, the period $T_\delta(E)$ tends to $T(E)$ along with derivatives with respect to E (because the return time to $S_{\delta,1}^-$ tends to the billiard return time with derivatives with respect to the initial conditions).

For every $\delta > 0$, for each impact point M^j , $j = 1, \dots, k^*$, let us choose the region $\mathcal{U}_{\delta,j}$ as in Lemma 4.3 (i.e., this region consists of the orbits of (21) that connect the cross-sections $S_{\delta,j}^-$ and $S_{\delta,j}^+$). Since L^* is regular periodic orbit, its impacts are distinct, so the regions $\mathcal{U}_{\delta,j}$ do not overlap for different j . Let \mathcal{U}^0 be a sufficiently small, yet independent of δ , open neighborhood of the part of L_δ^* which is not covered by the union of $\mathcal{U}_{\delta,j}$.

We use (q, p) as coordinates in \mathcal{U}^0 ; note that for the points $(q, p) \in \mathcal{U}^0$, the q -component is bounded away from the billiard boundary. In $\mathcal{U}_{\delta,j}$, we use the flow-box coordinates $(\tau, \mathcal{E}, \mathcal{P})_j$ given by Lemma 4.3. We restrict the freedom in the choice of the flow-box coordinates by the requirement that $(\mathcal{E}, \mathcal{P})_j = 0$ on the periodic orbit L_δ^* .

Thus, we have covered a neighborhood of L_δ^* by a system of coordinate charts for all $\delta > 0$. The overlap region between \mathcal{U}^0 and $\mathcal{U}_{\delta,j}$ near the cross-sections $S_{\delta,j}^-$ and $S_{\delta,j}^+$ stays bounded away from the billiard boundary. Therefore, by Lemma 4.3 the transition map between the (q, p) -coordinates in \mathcal{U}^0 and the flow-box coordinates in $\mathcal{U}_{\delta,j}$ in the overlap region is symplectic, depends continuously on δ , and has a well-defined limit, in C^∞ , as $\delta \rightarrow 0$.

In other words, we have introduced the structure of a smooth symplectic manifold in the neighborhood of L_δ^* for $\delta > 0$, and this structure has a regular limit at $\delta = 0$. The equations of motion also have a regular limit in these coordinates: in \mathcal{U}^0 the equations converge to

$$\dot{q} = p, \quad \dot{p} = 0$$

at $\delta = 0$, and in $\mathcal{U}_{\delta,j}$ the equations are the same for all δ :

$$\dot{\tau} = 1, \quad \dot{\mathcal{E}} = 0, \quad \dot{\mathcal{P}} = 0.$$

The elliptic periodic orbit L_δ^* tends, as $\delta \rightarrow 0$, to the billiard periodic orbit $L^* = L_0^*$, which is a smooth curve in these coordinates (for every $\delta \geq 0$, the curve L_δ^* has the same equation $(\mathcal{E}, \mathcal{P}) = 0$ in the flow-box coordinates). As in Section 2.1, we can introduce action-angle coordinates near L_δ^* for all $\delta \geq 0$, with the resulting Hamiltonian as in (6), yet here the coefficients depend (continuously) on δ :

$$H_0(I_0, \theta, z; \delta) = \frac{1}{2} + \omega(\delta)I + \frac{1}{2}I^\top A(\delta)I + g(I_0, \theta, z; \delta), \quad (91)$$

where $g = g_0(I_0; \delta) + g_1(I_0; \delta)\hat{I} + O(\|z\|^4|I_0| + \|z\|^5)$ with $g_0 = O(I_0^3)$, $g_1 = O(I_0^2)$ (see notations after (6)). We stress that the symplectic transformation between the action-angle coordinates (I_0, θ, z) and the energy-time coordinates $(\tau, \mathcal{E}, \mathcal{P})$ defined in the near-impact regions, $\mathcal{U}_{\delta,j}$, is smooth for all $\delta \geq 0$.

Similar to Section 2.1, we denote the relation between the action-angle coordinates and the (q, p) -coordinates as $(q, p) = (\hat{q}(I_0, \theta, z), \hat{p}(I_0, \theta, z))$. Away from the billiard boundary, i.e., in \mathcal{U}^0 , this coordinate transformation is well-defined in the limit $\delta = 0$. However, near the impacts (in the regions $\mathcal{U}_{\delta,j}$) the relation between (q, p) and the action-angle coordinates acquires singularities at $\delta = 0$.

The periodic orbit L_δ^* corresponds to $(I_0, z) = 0$. We have $\dot{\theta} = \omega_0(\delta)$ on L_δ^* , i.e., $\theta = \omega_0(\delta)t$. Thus, in our notations, $L_\delta^* = (q_\delta^*(\theta), p_\delta^*(\theta)) = (\hat{q}(0, \theta, 0), \hat{p}(0, \theta, 0))$. On L_δ^* , equation (90) near the j -th impact point, where for $\delta = 0$ the impact occurs at θ_j , becomes

$$q_\delta^*(\theta) = M^j + \delta^{1/\alpha} \tilde{q}_{j,\delta} \left(\frac{\theta - \theta_j}{\omega_0(\delta)\delta^{1/\alpha}} \right), \quad (92)$$

where $\tilde{q}_{j,\delta}$ has all derivatives bounded, see Lemma 4.1. So, near impacts, the k -th derivative of $q_\delta^*(\theta)$ is of order $\delta^{-(k-1)/\alpha}$.

Like in Section 3, we will also perform the scaling $z \rightarrow \delta^{1/4}z$, $I_0 \rightarrow \delta^{1/2}I_0$. In the scaled variables, the motion in a small neighborhood of L_δ^* is described by the scaled single-particle Hamiltonian

$$H_{0,scal} = \delta^{-1/2} \left(H_0(\delta^{1/2}I_0, \theta, \delta^{1/4}z; \delta) - \frac{1}{2} \right) = \omega(\delta)I + \frac{1}{2}\delta^{1/2}I^\top A(\delta)I + O(\delta^{3/4}). \quad (93)$$

Note that away from the billiard boundary, we have

$$\hat{q}(\delta^{1/2}I_0, \theta, \delta^{1/4}z) - q_\delta^*(\theta) = O(\delta^{1/4}), \quad (94)$$

with all derivatives with respect to θ and the scaled variables z and I_0 . Note also that the scaling of I_0 and z induces the scaling $(\mathcal{E}, \mathcal{P}) \rightarrow (\delta^{1/2}\mathcal{E}, \delta^{1/4}\mathcal{P})$ for the flow-box coordinates near impacts; the scaled energy \mathcal{E} equals to $H_{0,scal}$ in (93). Recall that the symplectic transformation between the unscaled (I_0, θ, z) and $(\tau, \mathcal{E}, \mathcal{P})$ is smooth for all $\delta \geq 0$. By (93), scaling I_0 and \mathcal{E} by $\delta^{1/2}$ remains a smooth transformation in the limit $\delta = 0$. Similarly, since \mathcal{P} depends smoothly on $(\delta^{1/2}I_0, \delta^{1/4}z)$ and vanishes at the origin, it follows that its scaling by $\delta^{1/4}$ also remains smooth in this limit. Therefore, we conclude that the symplectic transformation between the scaled (I_0, θ, z) and the scaled $(\tau, \mathcal{E}, \mathcal{P})$ is smooth for all $\delta \geq 0$.

4.1.4 Conditions SP1-SP3 for the orbit L_δ^* .

Lemma 4.4. *Under Assumptions BD1 and BD2 on the billiard orbit, the periodic orbit L_δ^* satisfies Assumptions SP1-SP3 for all sufficiently small $\delta \geq 0$.*

Proof. For the family $L^*(E)$ of billiard periodic orbits that follow, with the speed $\|p\| = \sqrt{2E}$, the same path as L_0^* in the q -space, the period $T(E)$ equals to $\frac{L_0^*}{\sqrt{2E}}$. It decreases with the energy E . Since for small δ the period of $L_\delta^*(E)$ is close to the period of $L^*(E)$ with the derivatives with respect to E (see Section 4.1.3), the period of $L_\delta^*(E)$ also decreases with E , i.e., Assumption SP1 is verified.

By the closeness of the return maps near L_δ^* and L_0^* , the multipliers of L_δ^* are close to the multipliers of L_0^* . Hence, by assumption BD2, there are no low-order resonances for all small δ , i.e., assumption SP2 is satisfied.

The twist assumption SP3 for the orbit L_δ^* is an open condition on coefficients of system (91). Because of the continuity in δ , it is enough to check this condition at $\delta = 0$. We evaluate the necessary coefficients of (91) at $\delta = 0$ by analyzing the return map near $L_0^* = (I_0 = 0, z = 0)$.

The system of differential equations defined by the Hamiltonian (91) is, in restriction to the energy level $H_0 = E$ close to $E = \frac{1}{2}$, given by

$$\begin{aligned} \omega_0 I_0 &= E - \frac{1}{2} - \hat{\omega} \hat{I} + O(\hat{I}^2), \\ \dot{\theta} &= \omega_0 + (aI_0 + b\hat{I}) + O(I^2), \\ \dot{x}_j &= (\omega_j + b_j I_0 + (\hat{A}\hat{I})_j + O(I_0^2))y_j + O(\|z\|^3|I_0| + \|z\|^4), \\ \dot{y}_j &= -(\omega_j + b_j I_0 + (\hat{A}\hat{I})_j + O(I_0^2))x_j + O(\|z\|^3|I_0| + \|z\|^4) \quad (j = 1, \dots, d-1), \end{aligned} \quad (95)$$

where all the coefficients are taken at $\delta = 0$. Recall the notation: $I = (I_0, \hat{I})$, $z_j = (x_j, y_j)$, $I_j = \frac{1}{2}z_j^2$, $\omega = (\omega_0, \hat{\omega})$, $A = \begin{pmatrix} a & b \\ b^\top & \hat{A} \end{pmatrix}$, see Section 2.1.

For each value of E , the line $(x, y) = 0$ (equivalently $\hat{I} = 0$) corresponds to a periodic orbit. Its period equals to the travel time from $\theta = 0$ to $\theta = 2\pi$, i.e.,

$$T(E) = \frac{2\pi}{\omega_0 + a\frac{E-\frac{1}{2}}{\omega_0} + O((E-\frac{1}{2})^2)} = \frac{2\pi}{\omega_0} \left(1 - a\frac{E-\frac{1}{2}}{\omega_0^2} + O((E-\frac{1}{2})^2)\right)$$

The linear part of the system for (x_j, y_j) is rotation with the frequency $\omega_j(E) = \omega_j + b_j\frac{E-\frac{1}{2}}{\omega_0} + O((E-\frac{1}{2})^2)$. It follows that the linearization of the return map $(x, y) \mapsto (\bar{x}, \bar{y})$ at the fixed point $(x, y) = 0$ is the rotation of (x_j, y_j) , $j = 1, \dots, d-1$, by the angle

$$\Delta_j = T(E)\omega_j(E) = \frac{2\pi}{\omega_0} \left(\omega_j + \frac{E-\frac{1}{2}}{\omega_0} \left(b_j - \frac{\omega_j a}{\omega_0}\right) + O((E-\frac{1}{2})^2)\right). \quad (96)$$

System (91) at $\delta = 0$ describes the billiard flow, so its return map is the return map of a billiard flow. The billiard flow is invariant with respect to the energy and time scaling, in particular the return map at the energy level $E = \frac{1}{2}$ is conjugate to the return map at any E by the scaling $(q, p) \rightarrow (q, p/\sqrt{2E})$. The conjugacy implies that the rotation angles Δ_j in (96) are independent of E . Therefore,

$$b_j = \frac{\omega_j a}{\omega_0}, \quad j = 1, \dots, d-1. \quad (97)$$

Now, at $E = \frac{1}{2}$ we rewrite system (95) as

$$\begin{aligned} \dot{\theta} &= \omega_0 + O(\|z\|^4), \\ \dot{x}_j &= \left(\omega_j - \frac{b_j}{\omega_0}\hat{\omega}\hat{I} + (\hat{A}\hat{I})_j\right)y_j + O(\|z\|^4), \\ \dot{y}_j &= -\left(\omega_j - \frac{b_j}{\omega_0}\hat{\omega}\hat{I} + (\hat{A}\hat{I})_j\right)x_j + O(\|z\|^4) \quad (j = 1, \dots, d-1), \end{aligned}$$

The return map $(\theta = 0) \rightarrow (\theta = 2\pi)$ for this system coincides, up to $O(\|z\|^4)$ -terms with the (nonlinear) rotation of (x_j, y_j) , $j = 1, \dots, d-1$, to the angles

$$\frac{2\pi}{\omega_0} \left(\omega_j - \frac{b_j}{\omega_0}\hat{\omega}\hat{I} + (\hat{A}\hat{I})_j\right).$$

Thus, the map coincides, up to $O(\hat{I}^2)$ -terms, with the normal form (22) where the action-angle coordinates (I_j, Φ_j) are introduced such that $(x_j, y_j) = \sqrt{2I_j}(\cos(\Phi_j), \sin(\Phi_j))$, and the

matrix Ω is given by

$$\Omega_{jk} = \frac{2\pi}{\omega_0} \left(\hat{A}_{jk} - \frac{\omega_k b_j}{\omega_0} \right) = \frac{2\pi}{\omega_0} \left(\hat{A}_{jk} - \frac{a\omega_j \omega_k}{\omega_0^2} \right).$$

We have $\det \Omega \neq 0$ by Assumption BD2. This implies that conditions (8) and (9) of Assumption SP3 are satisfied. Indeed, using (97), we find

$$\det A = a \det \left(\hat{A} - \frac{1}{a} b^\top b \right) = a \det \left(\frac{\omega_0}{2\pi} \Omega \right) \neq 0,$$

and

$$\det A_\omega = \det \begin{pmatrix} 0 & \omega_0 & \hat{\omega} \\ \omega_0 & a & \frac{a}{\omega_0} \hat{\omega} \\ \hat{\omega}^\top & \frac{a}{\omega_0} \hat{\omega}^\top & \hat{A} \end{pmatrix} = \det \begin{pmatrix} -\frac{\omega_0^2}{a} & 0 & 0 \\ \omega_0 & a & \frac{a}{\omega_0} \hat{\omega} \\ \hat{\omega}^\top & \frac{a}{\omega_0} \hat{\omega}^\top & \hat{A} \end{pmatrix} = -\frac{\omega_0^2}{a} \det A \neq 0.$$

□

4.2 Multi-particle system

Now we analyze the multi-particle system (20) near the minimal line (26) of the averaged interaction potential U of (25). We have established that for all small $\delta \geq 0$ the single-particle Hamiltonian can be brought to the form (91). In these variables the multi-particle system is of the form

$$H = \sum_{n=1}^N \left[\omega(\delta) I^{(n)} + \frac{1}{2} I^{(n)\top} A(\delta) I^{(n)} + g(I_0^{(n)}, \theta^{(n)}, z^{(n)}; \delta) \right] + \delta \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W(q^{(n)} - q^{(m)}), \quad (98)$$

where, as $\delta \rightarrow 0$, the dependence $q^{(n)} = \hat{q}(I_0^{(n)}, \theta^{(n)}, z^{(n)})$ becomes singular near impacts, as described in Section 4.1. This singularity is quite mild in the flow-box coordinates $(\tau^{(n)}, \mathcal{E}^{(n)}, \mathcal{P}^{(n)})$: we have, uniformly for all $\delta \geq 0$, continuous dependence of $q^{(n)}$ on $(\tau^{(n)}, \mathcal{E}^{(n)}, \mathcal{P}^{(n)})$ and, for every fixed value of $\tau^{(n)}$, smooth dependence on $(\mathcal{E}^{(n)}, \mathcal{P}^{(n)})$, see Lemma 4.3.

Let us make the scaling $z^{(n)} \rightarrow \delta^{1/4} z^{(n)}$, $I_0^{(n)} \rightarrow \delta^{1/2} I_0^{(n)}$. By (93), we obtain the scaled

version of (98):

$$\begin{aligned}
H_{scal} &= \delta^{-1/2} H\left((\delta^{1/2} I_0^{(n)}, \theta^{(n)}, \delta^{1/4} z^{(n)})_{n=1, \dots, N}\right) \\
&= \sum_n \left[\omega(\delta) I^{(n)} + \delta^{1/2} \frac{1}{2} I^{(n)\top} A(\delta) I^{(n)} + O(\delta^{3/4}) \right] \\
&\quad + \delta^{1/2} \sum_{n \neq m} W(\hat{q}(\delta^{1/2} I_0^{(n)}, \theta^{(n)}, \delta^{1/4} z^{(n)}) - \hat{q}(\delta^{1/2} I_0^{(m)}, \theta^{(m)}, \delta^{1/4} z^{(m)})).
\end{aligned} \tag{99}$$

As we consider a neighborhood of the minimal line (26) which satisfies the collision-free Assumption IP2, we can assume that W is bounded with its derivatives (as in Theorem 2). Therefore, away from the impacts,

$$W(\hat{q}(I_0^{(n)}, \theta^{(n)}, z^{(n)}) - \hat{q}(I_0^{(m)}, \theta^{(m)}, z^{(m)})) = W(q_\delta^*(\theta^{(n)}; \delta) - q_\delta^*(\theta^{(m)}; \delta)) + O(\delta^{1/4}),$$

see (94).

Thus, when all the particles stay away from the billiard boundary, the Hamiltonian (99) is $O(\delta^{3/4})$ -close, with all derivatives, to the truncated Hamiltonian

$$H_{trun} = \sum_n \omega(\delta) I^{(n)} + \delta^{1/2} \sum_n \frac{1}{2} I^{(n)\top} A(\delta) I^{(n)} + \delta^{1/2} \sum_{n \neq m} W(q_\delta^*(\theta^{(n)}; \delta) - q_\delta^*(\theta^{(m)}; \delta)). \tag{100}$$

Note that we cannot claim the same near impacts, so there we need a more accurate comparison of systems (99) and (100).

We consider the Hamiltonians (99) and (100) close to the line

$$\mathcal{L}_{min}^* = \{(I_0^{(n)} = 0, \theta^{(n)} = \theta_{min}^{(n)} + c, z^{(n)} = 0)_{n=1, \dots, N}, c \in S^1\},$$

the phase space image of the minima line (26), i.e. the solution curve (3) in action angle coordinates). Take the codimension-one hypersurface

$$\Sigma_0 : \sum_{n=1}^N \theta^{(n)} = 0 \pmod{2\pi}. \tag{101}$$

By (28), this surface intersects the line \mathcal{L}_{min}^* at $\theta^{(n)} = \theta_{min}^{(n)}$. By (29), in the neighborhood of this intersection point, all $q^{(n)}$ are bounded away from the billiard boundary. As $\dot{\varphi} = \frac{1}{N} \sum_{n=1}^N \dot{\theta}^{(n)} = \omega_0(\delta) + O(\delta^{1/2}) \neq 0$, the hyper-surface (101) is transverse to the flows of both Hamiltonians (99) and (100). Now we show that despite the fact that we can guarantee the closeness of these Hamiltonians to each other only away from the billiard boundary, their Poincaré return maps to the cross-section (101) are sufficiently close to each other.

Lemma 4.5. *When Assumption IP3 (non-simultaneous impacts) is satisfied, the Poincaré return map of system (99) to the cross-section $\sum_{n=1}^N \theta^{(n)} = 0$ is $o(\delta^{1/2})$ -close, along with all derivatives, to the corresponding return map of the truncated system (100).*

Proof. We take several cross-sections $\Sigma_{c_k} : \frac{1}{N} \sum_{n=1}^N \theta^{(n)} = c_k$, $0 = c_1 < c_2 < \dots < c_K = 2\pi$, and consider the return map of the flow of (99) to Σ_0 as a composition of the maps from each consecutive cross-section to the next one (intermediate Poincaré maps). We show that each of these maps is well defined in the neighborhood of \mathcal{L}_{min}^* and is sufficiently close to the corresponding map for (100).

In a small neighborhood of \mathcal{L}_{min}^* , by the non-simultaneous impacts Assumption IP3, when $q^{(n)}$ is close to an impact with the billiard boundary for particle n , all the other particles stay away from the billiard boundary. Therefore, we can choose the cross-sections Σ_{c_k} in such a way that when the phase point travels between Σ_{c_k} and $\Sigma_{c_{k+1}}$, either all the particles are bounded away from the billiard boundary, or one particle gets close to a regular impact while all the other particles stay away from the boundary. In the first case, systems (99) and (100) are uniformly $O(\delta^{3/4})$ -close in C^r for any r in the region of the phase space between these cross-sections. Since the flight time between the cross-sections is bounded, this implies that the corresponding intermediate Poincaré maps are also $O(\delta^{3/4})$ -close, as required.

It remains to consider the second case, when one of the particles gets close to an impact. Let it be particle n_0 . Let the cross-sections before and after the impact be Σ_{c_k} and $\Sigma_{c_{k+1}}$, and let the value of $q^{(n_0)}$ on these cross-sections stay at a distance of order $\eta > 0$ from the impact for all small δ . Let us show that the map $\Sigma_{c_k} \rightarrow \Sigma_{c_{k+1}}$ by system (99) is $O(\eta\delta^{1/2})$ close, in C^r for any r , to the corresponding map by the system

$$\hat{H}_{scal} = \sum_n \left[\omega(\delta)I^{(n)} + \delta^{1/2} \frac{1}{2} I^{(n)\top} A(\delta)I^{(n)} + O(\delta^{3/4}) \right] + \delta^{1/2} \sum_{\substack{n,m \neq n_0 \\ n \neq m}} W(\hat{q}(\delta^{1/2}I_0^{(n)}, \theta^{(n)}, \delta^{1/4}z^{(n)}) - \hat{q}(\delta^{1/2}I_0^{(m)}, \theta^{(m)}, \delta^{1/4}z^{(m)})) \quad (102)$$

(i.e., we switched off the interaction with the particle n_0), and, similarly, the map $\Sigma_{c_k} \rightarrow \Sigma_{c_{k+1}}$ by system (100) is $O(\eta\delta^{1/2})$ -close, in C^r for any r , to the corresponding map by the system

$$\hat{H}_{trun} = \sum_n \left[\omega(\delta)I^{(n)} + \delta^{1/2} \frac{1}{2} I^{(n)\top} A(\delta)I^{(n)} \right] + \delta^{1/2} \sum_{\substack{n,m \neq n_0 \\ n \neq m}} W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)})). \quad (103)$$

Indeed, since the particle n_0 between Σ_{c_k} and $\Sigma_{c_{k+1}}$ is close to impact, we can use the scaled flow-box coordinates $(\tau^{(n_0)}, \mathcal{E}^{(n_0)}, \mathcal{P}^{(n_0)})$ for the n_0 -th particle. As explained at the

end of section 4.1.3, the transformation between the scaled $(\tau^{(n_0)}, \mathcal{E}^{(n_0)}, \mathcal{P}^{(n_0)})$ coordinates and the scaled $(\theta^{(n_0)}, I_0^{(n_0)}, z^{(n_0)})$ is non-singular in the limit $\delta = 0$. We use the notation $(\tau = \tau^{(n_0)}, \mathcal{E} = \mathcal{E}^{(n_0)}, \mathcal{P} = \mathcal{P}^{(n_0)}, X = \{(I_0^{(n)}, \theta^{(n)}, z^{(n)})\}_{n \neq n_0})$ in the layer between Σ_{c_k} and $\Sigma_{c_{k+1}}$. The Hamiltonian (99) in these scaled coordinates can be schematically written as

$$H_{scal} = \mathcal{E} + H_1(X; \delta) + \delta^{1/2} H_2(\tau, \delta^{1/2} \mathcal{E}, \delta^{1/4} \mathcal{P}, X; \delta), \quad (104)$$

and the Hamiltonian (102) is given by

$$\hat{H}_{scal} = \mathcal{E} + H_1(X; \delta). \quad (105)$$

Here H_1 is the Hamiltonian of the subsystem corresponding to all the particles but n_0 :

$$H_1 = \sum_{n \neq n_0} \left[\omega(\delta) I^{(n)} + \delta^{1/2} \frac{1}{2} I^{(n)\top} A(\delta) I^{(n)} + O(\delta^{3/4}) \right] \\ + \delta^{1/2} \sum_{\substack{n, m \neq n_0 \\ n \neq m}} W(\hat{q}(\delta^{1/2} I_0^{(n)}, \theta^{(n)}, \delta^{1/4} z^{(n)}) - \hat{q}(\delta^{1/2} I_0^{(m)}, \theta^{(m)}, \delta^{1/4} z^{(m)})).$$

Since these particles stay away of impacts, H_1 is C^r for any fixed r , uniformly for all small $\delta \geq 0$. The term H_2 describes the interaction of the particle n_0 with the rest of the particles⁶:

$$H_2 = 2 \sum_{m \neq n_0} W(q(\tau, \delta^{1/2} \mathcal{E}, \delta^{1/4} \mathcal{P}) - \hat{q}(\delta^{1/2} I_0^{(m)}, \theta^{(m)}, \delta^{1/4} z^{(m)}))$$

where we denote $q(\tau, \delta^{1/2} \mathcal{E}, \delta^{1/4} \mathcal{P}) = \hat{q}(\delta^{1/2} I_0^{(n_0)}, \theta^{(n_0)}, \delta^{1/4} z^{(n_0)})$. By Lemma 4.3, H_2 is a C^∞ function of $(\delta^{1/2} \mathcal{E}, \delta^{1/4} \mathcal{P}, X)$ and depends continuously, in C^∞ , on τ and $\delta \geq 0$. Moreover, as long as $q^{(n_0)}$ stays away from the billiard boundary, H_2 depends smoothly on τ as well, for all $\delta \geq 0$.

Since Σ_{c_k} and $\Sigma_{c_{k+1}}$ stay bounded away from the impact, one obtains that near these cross-sections

$$\frac{d}{dt} \tau = \partial_{\mathcal{E}} H_{scal} = 1 + O(\delta)$$

uniformly for all $\delta \geq 0$, i.e., it stays bounded away from zero. Hence, hypersurfaces of constant τ are cross-sections to the flow of H_{scal} for all $\delta \geq 0$; the same is true for the flow of \hat{H}_{scal} . Thus, we choose constant- τ cross-sections, Σ_k^τ and Σ_{k+1}^τ close to Σ_{c_k} and, respectively, $\Sigma_{c_{k+1}}$. Because all these cross-sections are bounded away from impacts, the flows of H_{scal} and \hat{H}_{scal} are $O(\delta^{1/2})$ -close in C^r in between Σ_{c_k} and Σ_k^τ and in between Σ_{k+1}^τ and $\Sigma_{c_{k+1}}$,

⁶Recall that for the convenience of notation each term $W(q^{(n)} - q^{(m)})$ appears twice in the double sum in (98).

for any r . Since the flight time from Σ_{c_k} and $\Sigma_{c_{k+1}}$ is $O(\eta)$, the flight times between the cross-sections Σ_{c_k} and Σ_k^τ and between Σ_{k+1}^τ and $\Sigma_{c_{k+1}}$ are also bounded by $O(\eta)$, so the corresponding maps for H_{scal} and \hat{H}_{scal} between these cross-sections are $O(\eta\delta^{1/2})$ -close.

We are left to show that the maps from Σ_k^τ to Σ_{k+1}^τ for system (104) (H_{scal}) and system (105) (\hat{H}_{scal}) on any given energy level are $O(\eta\delta^{1/2})$ -close in C^r for any r . The equations of motion defined by (104) on the energy level $H_{scal} = E$ are given by

$$\begin{aligned}\dot{\tau} &= 1 + \delta^{1/2}\partial_{\mathcal{E}}H_2, \\ \dot{\mathcal{P}} &= \delta^{1/2}\{\mathcal{P}, H_2\}, \\ \dot{X} &= \{X, H_1\} + \delta^{1/2}\{X, H_2\},\end{aligned}$$

where $\{\cdot, \cdot\}$ are the Poisson brackets in the (\mathcal{P}, X) space. To have the equations in a closed form, we substitute the energy \mathcal{E} of particle (n_0) by its expression which can be found from (104) by the implicit function theorem:

$$\mathcal{E} = E - H_1(X; \delta) - \delta^{1/2}H_2(\tau, \delta^{1/2}\mathcal{E}, \delta^{1/4}\mathcal{P}, X; \delta) = E - H_1(X; \delta) - \delta^{1/2}H_3(\tau, \delta^{1/4}\mathcal{P}, X; \delta),$$

where H_3 is continuous in τ and C^∞ -smooth in $\delta^{1/4}\mathcal{P}, X$ uniformly for all τ and $\delta \geq 0$.

Choosing τ as the time variable, we obtain a non-autonomous system of the form

$$\frac{d}{d\tau}\dot{\mathcal{P}} = O(\delta^{3/4}), \quad \frac{d}{d\tau}X = \{X, H_1\} + O(\delta^{1/2}),$$

where the right-hand sides are continuous in the new time τ and C^∞ in \mathcal{P}, X for all small $\delta \geq 0$. The right-hand side of this system is $O(\delta^{1/2})$ -close with all derivatives with respect to \mathcal{P}, X , to

$$\frac{d}{d\tau}\dot{\mathcal{P}} = 0, \quad \frac{d}{d\tau}X = \{X, H_1\},$$

which is the system (105) restricted to any constant energy level. The constant- $\Delta\tau$ maps of these systems are therefore $O(\Delta\tau\delta^{1/2})$ -close in C^r for any r . This is what we need since $\Delta\tau$, the flight time from Σ_k^τ to Σ_{k+1}^τ , is $O(\eta)$.

This proves that the Poincaré map $\Sigma_{c_k} \rightarrow \Sigma_{c_{k+1}}$ for the Hamiltonian H_{scal} of (99) is $O(\eta\delta^{1/2})$ -close, in C^r for any r , to the corresponding map for the Hamiltonian \hat{H}_{scal} of (102).

Next we show that the corresponding maps for systems (100) and (103) are also $O(\eta\delta^{1/2})$ -close in C^r for any r . Here, we do not use the flow-box coordinates and instead of showing the $O(\eta\delta^{1/2})$ -closeness of the Poincaré maps between constant- τ cross-sections we show the $O(\eta\delta^{1/2})$ -closeness of the Poincaré maps between constant- $\theta^{(n_0)}$ cross-sections, Σ_k^θ before impact and Σ_{k+1}^θ after impact. Recall that only the particle n_0 is near impact for the flow between $\Sigma_{c_k} \rightarrow \Sigma_{c_{k+1}}$. Hence the flows of (100) and (103) between Σ_{c_k} and Σ_k^θ and

between Σ_{k+1}^θ and $\Sigma_{c_{k+1}}$ are $O(\delta^{1/2})$ -close, with derivatives, for all $\delta \geq 0$. Since the flight time between these cross-sections is bounded by η , the difference between the corresponding Poincaré maps for systems (100) and (103) is $O(\eta\delta^{1/2})$ small. So we need to examine only the flow between Σ_k^θ and Σ_{k+1}^θ . As explained in Section 4.1.3, the terms $W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)}))$ in H_{trun} of (100) and \hat{H}_{trun} of (103) are C^∞ functions of all the variables if $n, m \neq n_0$. The terms $W(q_\delta^*(\theta^{(n_0)}) - q_\delta^*(\theta^{(m)}))$ with $m \neq n_0$ in (100) are C^∞ functions of $\theta^{(m)}$ uniformly for all $\theta^{(n_0)}$ and $\delta \geq 0$. Restricting the flow to a constant energy level and choosing $\theta^{(n_0)}$ as a new time variable, the right-hand sides of the corresponding non-autonomous systems obtained from (100) and (103) are $O(\delta^{1/2})$ -close with derivatives with all the variables but $\theta^{(n_0)}$. This implies the $O(\eta\delta^{1/2})$ -closeness of their constant- $\theta^{(n_0)}$ maps, which implies the required $O(\eta\delta^{1/2})$ -closeness of the maps $\Sigma_{c_k} \rightarrow \Sigma_{c_{k+1}}$ for the Hamiltonians H_{trun} and \hat{H}_{trun} .

Because there is no dependence on the q -coordinates of the particle n_0 in equations (102) and (103), and this is the only source of singularity at $\delta = 0$ (all the other particles are away from the impacts), the Hamiltonians \hat{H}_{scal} and \hat{H}_{trun} are $O(\delta^{3/4})$ -close in C^r for any r . Therefore, their Poincaré maps are also $O(\delta^{3/4})$ -close. This, finally, proves that the maps $\Sigma_{c_k} \rightarrow \Sigma_{c_{k+1}}$ by system (99) and system (100) are $O(\eta\delta^{1/2})$ -close.

Taking the composition of the Poincaré maps from each consecutive cross-section to the next one, we obtain that the return maps to Σ_0 for systems (99) and (100) are $O(\eta\delta^{1/2})$ -close, in C^r for any r . Since η can be taken as small as we want, we can also allow it to go sufficiently slowly to zero as $\delta \rightarrow 0$, thus completing the proof of the lemma. \square

We now evaluate the Poincaré map for the truncated system (100). We average the truncated Hamiltonian H_{trun} with respect to the motion along the periodic orbit L_δ^* to establish an analogue of Lemma 3.1, but, due to the loss of smoothness at impacts, the proof is different – we utilize explicitly the pairwise structure of the interaction potential instead of using the Fourier expansion of Lemma 3.1.

First, we introduce the following terminology:

Definition 4.6. A function G of $\left((I_0^{(1)}, \theta^{(1)}, z^{(1)}), \dots, (I_0^{(N)}, \theta^{(N)}, z^{(N)}) \right)$ is called *weakly regular* if it is C^∞ at $\delta > 0$ and satisfies the following conditions (i) and (ii) in the limit $\delta \rightarrow 0$:

(i) When all $q^{(n)}(I_0^{(n)}, \theta^{(n)}, z^{(n)})$ stay bounded away from the billiard boundary, the function G is of class C^∞ for all $\delta \geq 0$ and depends continuously, in C^∞ , on $\delta \geq 0$;

(ii) When exactly one of the particles is in the billiard boundary layer, i.e., for some n_0 , $q^{(n_0)}(I_0^{(n_0)}, \theta^{(n_0)}, z^{(n_0)})$ is near an impact point, the function G is C^∞ with respect to $(I^{(n_0)}, z^{(n_0)})$ and $(\{I_0^{(n)}, \theta^{(n)}, z^{(n)}\}_{n \neq n_0})$ for each fixed $\theta^{(n_0)}$ and all $\delta \geq 0$, and it depends continuously, as a C^∞ function of $(I^{(n_0)}, z^{(n_0)})$ and $(\{I_0^{(n)}, \theta^{(n)}, z^{(n)}\}_{n \neq n_0})$, on $\theta^{(n_0)}$ and $\delta \geq 0$.

In particular, since we showed that if the particle n_0 is close to impact $q_\delta^*(\theta^{(n_0)})$ depends

continuously on $\theta^{(n_0)}$, the functions $W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)}))$ in (100) are weakly regular. Note that we do not consider the case when there are more than one particle in the boundary layer because it is not possible by the non-simultaneous impacts assumption IP3.

Lemma 4.7. *There exists a weakly regular, symplectic change of coordinates which brings the truncated Hamiltonian (100) to the form*

$$H = \sum_{n=1}^N \omega(\delta) I^{(n)} + \delta^{1/2} \sum_{n=1}^N \frac{1}{2} I^{(n)} A(\delta) I^{(n)} + \delta^{1/2} U(\theta^{(1)}, \dots, \theta^{(N)}; \delta) + \delta^{3/4} G. \quad (106)$$

Here G is a weakly regular function, and the averaged potential is given by:

$$U(\theta^{(1)}, \dots, \theta^{(N)}; \delta) = \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W_{avg}(\theta^{(n)} - \theta^{(m)}; \delta), \quad (107)$$

where

$$W_{avg}(\theta^{(n)} - \theta^{(m)}; \delta) = \frac{1}{2\pi} \int_0^{2\pi} W(q_\delta^*(s + \theta^{(n)}) - q_\delta^*(s + \theta^{(m)})) ds. \quad (108)$$

All the derivatives of W_{avg} with respect to θ depend continuously on δ for small $\delta \geq 0$.

Proof. Let $(u, v) = (\frac{1}{2}(\theta^{(n)} + \theta^{(m)}), \frac{1}{2}(\theta^{(n)} - \theta^{(m)}))$. Define

$$\Psi_0(u, v; u_0, \delta) = \frac{1}{\omega_0(\delta)} \left[\int_{u_0}^u W(q_\delta^*(s + v) - q_\delta^*(s - v)) ds - (u - u_0) W_{avg}(2v; \delta) \right], \quad (109)$$

so $\Psi_0(u, -v; u_0, \delta) = \Psi_0(u, v; u_0, \delta)$. It follows from (108) and the 2π -periodicity of q_δ^* that

$$\Psi_0(u + 2\pi, v; u_0, \delta) = \Psi_0(u, v; u_0, \delta). \quad (110)$$

Let

$$\Psi(\theta^{(1)}, \dots, \theta^{(N)}; u_0, \delta) = \sum_{m_1 \neq m_2} \Psi_0\left(\frac{1}{2}(\theta^{(m_1)} + \theta^{(m_2)}), \frac{1}{2}(\theta^{(m_1)} - \theta^{(m_2)}); u_0, \delta\right). \quad (111)$$

For positive δ , it is identical to the function (48) of Section 3. Consider the symplectic coordinate change (its smoothness properties in the limit $\delta \rightarrow 0$ are discussed below)

$$I_0^{(n)} \rightarrow I_0^{(n)} - \delta^{1/2} \partial_{\theta^{(n)}} \Psi, \quad n = 1, \dots, N. \quad (112)$$

By (110), the right-hand side does not change when we add 2π to all phases $\theta^{(1)}, \dots, \theta^{(N)}$, i.e., this is indeed a well-defined coordinate transformation in a neighborhood of \mathcal{L}_{min}^* . Dif-

ferentiating (109), we obtain

$$\omega_0(\delta)\partial_u\Psi = \omega_0(\delta)\partial_{\theta^{(n)}}\Psi + \omega_0(\delta)\partial_{\theta^{(m)}}\Psi = W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)})) - W_{avg}(\theta^{(n)} - \theta^{(m)}; \delta).$$

As $\partial_{\theta^{(n)}}\Psi = \sum_{m \neq n} \partial_u \Psi_0(\frac{1}{2}(\theta^{(n)} + \theta^{(m)}), \frac{1}{2}(\theta^{(n)} - \theta^{(m)}); u_0, \delta)$, the transformation (112) makes

$$\sum_{n=1}^N \omega_0(\delta)I_0^{(n)} \rightarrow \sum_{n=1}^N \omega_0(\delta)(I_0^{(n)} - \delta^{1/2} \sum_{m \neq n} \partial_u \Psi_0(\frac{1}{2}(\theta^{(n)} + \theta^{(m)}), \frac{1}{2}(\theta^{(n)} - \theta^{(m)}); u_0, \delta))$$

so it brings the Hamiltonian (100) to the required form (106).

We need to show that one can choose u_0 such that the expression

$$\partial_{\theta^{(n)}}\Psi_0(\frac{1}{2}(\theta^{(n)} + \theta^{(m)}), \frac{1}{2}(\theta^{(n)} - \theta^{(m)}); u_0, \delta)$$

in the formula for the coordinate transformation (112) is weakly regular for all n and m . Let

$$\Psi_1(u, v; u_0, \delta) = \int_{u_0}^u W(q_\delta^*(s+v) - q_\delta^*(s-v))ds.$$

Lemma 4.8. *For u_0 chosen such that $(u_0 \pm \frac{1}{2}(\theta_{min}^{(n)} - \theta_{min}^{(m)}))/\omega_0$ are bounded away from the impact moments t^j for all m, n , and j , the function $\Psi_1(\frac{\theta^{(n)} + \theta^{(m)}}{2}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_0, \delta)$ and its derivatives with respect to $\theta^{(n)}, \theta^{(m)}$ are weakly regular.*

Proof. Let us divide the interval of integration to subintervals $[u_j, u_{j+1}]$, $j = 0, \dots, J$ such that the following holds: the last interval is $[u_J, u_{J+1} = \frac{\theta^{(n)} + \theta^{(m)}}{2}]$; on each sub-interval at most one of the particles is in the boundary layer (i.e., for all s in each sub-interval, either $q_\delta^*(s + \frac{1}{2}(\theta^{(n)} - \theta^{(m)}))$, or $q_\delta^*(s - \frac{1}{2}(\theta^{(n)} - \theta^{(m)}))$, or both are outside the boundary layer); and all end points but the last one are at a finite distance from the impact points: $(u_j \pm \frac{1}{2}(\theta_{min}^{(n)} - \theta_{min}^{(m)}))/\omega_0$ are bounded away from the impact moments for $j = 0, \dots, J$.

Such choice of sufficiently small intervals is possible: because we consider $\theta^{(n)}$ and $\theta^{(m)}$ close to $\theta_{min}^{(n)}$ and $\theta_{min}^{(m)}$, the non-simultaneous impacts assumption IP3 implies that $\theta^{(n)} - \theta^{(m)}$ can not get close to $\omega_0(t^j - t^k)$ for $n \neq m$, hence it is impossible that $s_1 + \frac{1}{2}(\theta^{(n)} - \theta^{(m)})$ and $s_2 - \frac{1}{2}(\theta^{(n)} - \theta^{(m)})$ get simultaneously close to the impact phases $\omega_0 t^j$ and $\omega_0 t^k$ if s_1 and s_2 belong to the same small interval.

By construction,

$$\Psi_1(\frac{\theta^{(n)} + \theta^{(m)}}{2}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_0, \delta) = \sum_{j=0}^J \Psi_1(u_{j+1}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_j, \delta). \quad (113)$$

For all $j < J$, we show that $\Psi_1(u_{j+1}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_j, \delta)$ is C^∞ for all small $\delta \geq 0$ (hence it is weakly regular, and its derivatives are weakly regular as well). Indeed, if both $q_\delta^*(s \pm \frac{\theta^{(n)} - \theta^{(m)}}{2})$ are not in the boundary layer, for all s from the integration interval $[u_j, u_{j+1}]$, then the integrand $W(q_\delta^*(s + \frac{\theta^{(n)} - \theta^{(m)}}{2}) - q_\delta^*(s - \frac{\theta^{(n)} - \theta^{(m)}}{2}); \delta)$ is a C^∞ function for all $\delta \geq 0$ and the claim follows. If, say, the first term, $q_\delta^*(s + \frac{\theta^{(n)} - \theta^{(m)}}{2})$ is in the boundary layer for some value of $s \in (u_j, u_{j+1})$, then we shift the integration interval by $\frac{1}{2}(\theta^{(n)} - \theta^{(m)})$ to establish

$$\Psi_1(u_{j+1}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_j, \delta) = \int_{u_j + \frac{1}{2}(\theta^{(n)} - \theta^{(m)})}^{u_{j+1} + \frac{1}{2}(\theta^{(n)} - \theta^{(m)})} W(q_\delta^*(s') - q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)}))) ds'. \quad (114)$$

The term $q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)}))$ is away from the billiard boundary for the integration interval, $q_\delta^*(s')$ is away from the boundary layer at the limits of integration. Therefore, the integrand $W(q_\delta^*(s') - q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)})))$ is a C^∞ -function of $\theta^{(n)}$ and $\theta^{(m)}$, uniformly continuous in C^∞ for all s' and all $\delta \geq 0$. Moreover, it is C^∞ in s also near the limits of integration. It immediately follows that Ψ_1 given by (114) is C^∞ as required. Similarly, if the second term, $q_\delta^*(s - \frac{\theta^{(n)} - \theta^{(m)}}{2})$ is near an impact, we shift the integration interval by $-\frac{1}{2}(\theta^{(n)} - \theta^{(m)})$ and establish the same smoothness result.

For the last segment, since $u_{J+1} = \frac{\theta^{(n)} + \theta^{(m)}}{2}$, and $v = \frac{\theta^{(n)} - \theta^{(m)}}{2}$, if both $\theta^{(n)}/\omega_0$ and $\theta^{(m)}/\omega_0$ are bounded away from the impact moments, the same arguments as above show that $\Psi_1(\frac{\theta^{(n)} + \theta^{(m)}}{2}, \frac{1}{2}(\theta^{(n)} - \theta^{(m)}); u_J, \delta)$ are smooth as required. On the other hand, if, say, the particle n is in the boundary layer (and hence the particle m is not in the boundary layer) we write the integral in the form of (114):

$$\Psi_1(u_{J+1}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_J, \delta) = \int_{u_J + \frac{1}{2}(\theta^{(n)} - \theta^{(m)})}^{\theta^{(n)}} W(q_\delta^*(s') - q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)}))) ds'. \quad (115)$$

As above, we have that $W(q_\delta^*(s') - q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)})))$ is a C^∞ -function of $\theta^{(n)}$ and $\theta^{(m)}$, uniformly continuous in C^∞ for all s' and all $\delta \geq 0$. Moreover, it is C^∞ in s also near the lower limit of integration. The upper limit of integration does not depend on $\theta^{(m)}$, so we conclude that the integral is C^∞ with respect to $\theta^{(m)}$, i.e., Ψ_1 and $\partial_{\theta^{(m)}} \Psi_1$ given by (115) are weakly regular, and

$$\begin{aligned} \partial_{\theta^{(n)}} \Psi_1(u_{J+1}, \frac{\theta^{(n)} - \theta^{(m)}}{2}; u_J, \delta) &= W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)})) \\ &\quad - \frac{1}{2} W(q_\delta^*(u_J + \frac{1}{2}(\theta^{(n)} - \theta^{(m)})) - q_\delta^*(u_J - \frac{1}{2}(\theta^{(n)} - \theta^{(m)}))) \\ &\quad - \int_{u_J + \frac{1}{2}(\theta^{(n)} - \theta^{(m)})}^{\theta^{(n)}} W'(q_\delta^*(s') - q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)}))) \partial_{\theta^{(n)}} q_\delta^*(s' - (\theta^{(n)} - \theta^{(m)})) ds' \end{aligned}$$

is also weakly regular since $s' - (\theta^{(n)} - \theta^{(m)})$ is away from impact along this integration

interval. The similar weakly regularity results hold true if the particle m , and not the particle n , is in the boundary layer.

Thus, all the terms in (113) are weakly regular, along with the derivatives with respect to $\theta^{(n)}$ and $\theta^{(m)}$, which gives the lemma. \square

Note that by the periodicity of q_δ^* ,

$$W_{avg}(\theta^{(n)} - \theta^{(m)}; \delta) = \frac{1}{2\pi} \Psi_1(u_0 + 2\pi, u_0, \frac{\theta^{(n)} - \theta^{(m)}}{2}; \delta), \quad (116)$$

and so the right-hand side does not depend on u_0 . In particular we can choose u_0 as needed for Lemma 4.8, so the average potential W_{avg} is weakly regular. Since we can always shift both $\theta^{(n)}$ and $\theta^{(m)}$ away from the impacts, the weak regularity of the average potential W_{avg} means that it is C^∞ for all small $\delta \geq 0$.

Now, by Lemma 4.8, we obtain that since

$$\Psi_0(u, v; u_0, \delta) = \frac{1}{\omega_0} \Psi_1(u, u_0, v; \delta) - \frac{u - u_0}{\omega_0} W_{avg}(2v; \delta),$$

the transformation (112) is also weakly regular. \square

Like in Lemma 4.5, omitting the weakly regular $O(\delta^{3/4})$ -term in the Hamiltonian (106) results only in $O(\delta^{3/4})$ corrections to the return map to an interior cross-section near \mathcal{L}_{min}^* . Thus, the return map for the truncated averaged Hamiltonian

$$H(I, \theta) = \sum_{n=1}^N \omega(\delta) I^{(n)} + \delta^{1/2} \sum_{n=1}^N \frac{1}{2} I^{(n)} A(\delta) I^{(n)} + \delta^{1/2} U(\theta^{(1)}, \dots, \theta^{(N)}; \delta) \quad (117)$$

near \mathcal{L}_{min}^* is $o(\delta^{1/2})$ -close the return map for the scaled Hamiltonian (99).

This Hamiltonian has the same form as the truncated Hamiltonian of (46). The parameters $\omega(\delta)$, $A(\delta)$, and the averaged potential $U(\theta^{(1)}, \dots, \theta^{(N)}; \delta)$ depend continuously on δ and satisfy, for all small $\delta \geq 0$, the non-degeneracy assumptions as in Theorem 1. So we finish the proof of Theorem 3 in the same way as in Theorem 1. Namely, applying Lemma 3.2, we find that the Poincaré return map for system (117) is $O(\delta^{3/4})$ -close, with all derivatives, to the time- $\frac{2\pi}{\omega_0}$ map for system (55). Hence, the return map for (99) is $o(\delta^{1/2})$ -close to the time- $\frac{2\pi}{\omega_0}$ map for system (55). Now, the same arguments as in Lemma 3.3 show that if we make $O(\delta^{-1/2})$ iterations of the Poincaré map of (99), the result is $o(1)_{\delta \rightarrow 0}$ -close to the time-1 map for the Hamiltonian (58). The latter map has, by Lemma 3.4, a positive measure set of invariant KAM tori, hence so does the rescaled system (99), as well as the original system (98).

5 Gas in a rectangular box

Here we study the motion near the family of fast vertical periodic orbits in a box, proving Theorem 4 (the non-simultaneous impacts case, see Sections 5.1, 5.2 and 5.3) and Theorem 5 (the simultaneous impacts case, Sections 5.1, 5.3 and 5.4).

5.1 Action-angle coordinates near vertical periodic orbits.

Since the single-particle degrees of freedom decouple, and only the vertical direction dynamics are fast and billiard-like, the single-particle theory developed in Section 4.1 applies to the one-dimensional vertical motion, where $q = q_d \in R^1$ and the billiard corresponds to a particle bouncing between the end points of the interval $[0, \pi]$. At positive δ , the vertical motion of a single particle is described by the Hamiltonian

$$H_d = \frac{p^2}{2} + \delta V_d(q) = \frac{p^2}{2} + \frac{\delta}{Q_d(q)^\alpha}, \quad (118)$$

see (30), (31). One can introduce action-angle variables (I, θ) for this system. The periodic orbit that lies in the energy level $H = \frac{1}{2}$ is denoted as $q = q_\delta^*(\theta)$; it tends to the saw-tooth (36) as $\delta \rightarrow 0$. The vertical action I is a one-to-one, smooth function of the vertical energy and is a constant of motion. Thus, the single-particle vertical motion is governed by

$$H_d(I; \delta) = \frac{(\hat{p}_d(I, \theta; \delta))^2}{2} + \delta V_d(\hat{q}_d(I, \theta; \delta)) - \frac{1}{2} = \omega_0(\delta)I + \frac{1}{2}a(\delta)I^2 + O(I^3), \quad I \in \mathbb{R}^1. \quad (119)$$

This is analogous to formula (91) of Section 4.1 but, contrary to the multidimensional single-particle theory, there are no z -variables nor the corresponding \mathcal{P} -variables.

Using the action-angle coordinates in the vertical direction, the multi-particle Hamiltonian (33) takes the following form:

$$H = \sum_{n=1}^N H_d(I^{(n)}; \delta) + \delta \sum_{n=1, \dots, N} \sum_{i=1}^{d-1} \left[\frac{(p_{\xi, i}^{(n)})^2}{2} + V_i(\xi_i^{(n)}) \right] + \delta \sum_{\substack{n, m=1, \dots, N \\ n \neq m}} W(\hat{q}_d(I^{(n)}, \theta^{(n)}; \delta) - \hat{q}_d(I^{(m)}, \theta^{(m)}; \delta), \xi^{(n)} - \xi^{(m)}); \quad (120)$$

recall that (ξ, p_ξ) denote the non-vertical coordinates: $\xi_i^{(n)} = q_i^{(n)}$, $p_{\xi, i}^{(n)} = p_i^{(n)}$, $i = 1, \dots, d-1$, $n = 1, \dots, N$.

Scaling the vertical actions as $\delta^{1/2}I$ and dividing the Hamiltonian by $\delta^{1/2}$, we obtain

$$H_{scal} = \sum_{n=1}^N \omega_0(\delta)I^{(n)} + \delta^{1/2} \sum_{n=1}^N \left(\frac{1}{2}a(\delta)(I^{(n)})^2 + \sum_{i=1}^{d-1} \left[\frac{(p_{\xi,i}^{(n)})^2}{2} + V_i(\xi_i^{(n)}) \right] \right) + O(\delta) \\ + \delta^{1/2} \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W(\hat{q}_d(\delta^{1/2}I^{(n)}, \theta^{(n)}; \delta) - \hat{q}_d(\delta^{1/2}I^{(m)}, \theta^{(m)}; \delta), \xi^{(n)} - \xi^{(m)}). \quad (121)$$

Below, we establish the existence of an elliptic periodic orbit for system (121) by comparing it with the averaged system defined by the Hamiltonian

$$H_{avg} = \sum_{n=1}^N \omega_0(\delta)I^{(n)} + \delta^{1/2} \sum_{n=1}^N \left(\frac{1}{2}a(\delta)(I^{(n)})^2 + \sum_{i=1}^{d-1} \left[\frac{(p_{\xi,i}^{(n)})^2}{2} + V_i(\xi_i^{(n)}) \right] \right) \\ + \delta^{1/2} \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}; \delta) + \delta \tilde{G}(\theta, I, \xi, p; \delta). \quad (122)$$

This is an analogue of the averaged system (106) (just the higher order terms are of order δ and not of order $\delta^{3/4}$ as in (106)). Here W_{avg} is the pairwise interaction potential averaged over the vertical oscillations in the system where the interaction between particles is switched off (see (120)):

$$H = \sum_{n=1}^N H_d(I^{(n)}; \delta) + \delta \sum_{n=1,\dots,N} \sum_{i=1}^{d-1} \left[\frac{(p_{\xi,i}^{(n)})^2}{2} + V_i(\xi_i^{(n)}) \right].$$

The vertical oscillations at $I^{(n)} = 0$ correspond to the choreographic solution (cf. (3)):

$$\mathbf{L}^*(\theta, \xi) = \{q_{1,\dots,d-1}^{(n)} = \xi_i^{(n)}, q_d^{(n)} = q_d^*(\omega_0 t + \theta^{(n)}), p_{1,\dots,d-1}^{(n)} = 0, p_d^{(n)} = p_d^*(\omega_0 t + \theta^{(n)})\}_{n=1}^N. \quad (123)$$

Thus,

$$W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}; \delta) = \frac{1}{2\pi} \int_0^{2\pi} W(q_\delta^*(s + \theta^{(n)}) - q_\delta^*(s + \theta^{(m)}), \xi^{(n)} - \xi^{(m)}) ds. \quad (124)$$

We denote the potential of the averaged system (122) by

$$U_\delta(\theta, \xi) = \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}; \delta) + \sum_{n=1}^N \sum_{i=1}^{d-1} V_i(\xi_i^{(n)}), \quad (125)$$

so $U_0(\theta, \xi) = U(\theta, \xi)$ of (25). The regularity properties of the potential U_δ and of the

correction term δG in (122) in the limit $\delta \rightarrow 0$, and their influence on the dynamics, are evaluated differently in the case of non-simultaneous and simultaneous impacts.

5.2 The non-simultaneous impacts case

The proof of Theorem 4 is, essentially, the same as for Theorem 3 - we just do not have here the fast variables z but, instead, have slow variables ξ (which, in fact, makes the situation simpler). As in Lemma 4.5 of Section 4.2, in the case where the impacts are non-simultaneous, the Poincaré map for the Hamiltonian (121) is $o(\delta^{1/2})$ -close to the Poincaré map for the truncated Hamiltonian H_{trun} :

$$H_{trun} = \sum_{n=1}^N \omega_0(\delta) I^{(n)} + \delta^{1/2} \sum_{n=1}^N \left(\frac{1}{2} a(\delta) (I^{(n)})^2 + \sum_{i=1}^{d-1} \left[\frac{(p_{\xi,i}^{(n)})^2}{2} + V_i(\xi_i^{(n)}) \right] \right) + \delta^{1/2} \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)}), \xi^{(n)} - \xi^{(m)}). \quad (126)$$

As in Lemma 4.7, we average the truncated Hamiltonian H_{trun} by performing the symplectic coordinate transformation defined by the $\delta^{1/2}$ -time map for the Hamiltonian Ψ defined as in (111), where Ψ_0 is defined as in (109), i.e.,

$$\Psi(\theta^{(1)}, \dots, \theta^{(N)}, \xi^{(1)}, \dots, \xi^{(N)}; u_0, \delta) = \sum_{m_1 \neq m_2} \Psi_0\left(\frac{1}{2}(\theta^{(m_1)} + \theta^{(m_2)}), \frac{1}{2}(\theta^{(m_1)} - \theta^{(m_2)}), \xi^{(m_1)} - \xi^{(m_2)}; u_0, \delta\right), \quad (127)$$

where

$$\Psi_0(u, v, w; u_0, \delta) = \frac{1}{\omega_0(\delta)} \left[\int_{u_0}^u W(q_\delta^*(s+v) - q_\delta^*(s-v), w; \delta) ds - (u - u_0) W_{avg}(2v, w; \delta) \right].$$

The only difference with Section 4.2 is that now Ψ also depends on the ξ variables. The same computations as in Lemma 4.7 show that the resulting near-identity transformation $(I^{(n)}, p_\xi^{(n)}) \rightarrow (I^{(n)}, p_\xi^{(n)}) - \delta^{1/2}(\partial_{\theta^{(n)}}, \partial_{\xi^{(n)}})\Psi$ is weakly regular and brings the truncated Hamiltonian (126) to the averaged form (122), where the error term δG is weakly regular (see Definition 4.6); the potential W_{avg} is a C^∞ function and depends continuously on δ , along with all derivatives, for all $\delta \geq 0$.

Like in Theorem 3, since no more than one particle can be near the box boundary at any given moment of time (this is the non-simultaneous impacts assumption), the weak regularity of G implies that omitting the δG terms in the averaged system (122) results only in $O(\delta)$ -corrections to the Poincaré return map. Thus, we obtain that the Poincaré return

map of the original system (121) is $o(\delta^{1/2})$ -close to the Poincaré return map for the system

$$H = \sum_{n=1}^N \omega_0(\delta) I^{(n)} + \delta^{1/2} \sum_{n=1}^N \left(\frac{1}{2} a(\delta) (I^{(n)})^2 + \sum_{i=1}^{d-1} \frac{(p_{\xi,i}^{(n)})^2}{2} \right) + \delta^{1/2} U_\delta(\theta, \xi) \quad (128)$$

where $U_\delta(\theta, \xi)$ is defined by (125). Since $U_\delta(\theta, \xi)$ and all its derivatives depend on δ continuously, the non-degeneracy assumption Box2 implies that the minimal line at $\delta = 0$ is non-degenerate and persists for small δ . So, we introduce local normal coordinates near this line:

$$(\theta^{(n)}, I^{(n)}) \rightarrow (\varphi, \psi, P, J),$$

in the same way as in (51),(52). By the translation invariance of U_δ , it is independent of φ :

$$U_\delta(\theta, \xi) = \hat{U}_\delta(\psi, \xi). \quad (129)$$

The potential \hat{U}_δ has a non-degenerate minimum at $(\psi = 0, \xi = \xi_{min})$. The Hamiltonian H of (128) becomes

$$H(P, \varphi, J, \psi, p_\xi, \xi; \delta) = \omega_0(\delta) P + \delta^{1/2} \left(\frac{1}{2N} a(\delta) P^2 + \frac{1}{2N} a(\delta) J^2 + \frac{p_\xi^2}{2} + \hat{U}_\delta(\psi, \xi) \right). \quad (130)$$

This Hamiltonian is similar to that in (53) but it is simpler, as there is only one fast degree of freedom (φ, P) . Moreover, H is independent of φ , so P is an integral which controls the period of the fast motion.

With this simplification in mind, let us follow the same procedure as is applied in Section 3 to Hamiltonian (53) and compute the Poincaré return map from $\varphi = 0$ to $\varphi = 2\pi$. Since the flight time depends only on P , the restriction of this map onto a fixed level of P is a constant-time map for the Hamiltonian (130) or, equivalently, an $O(\delta^{1/2})$ -time map for the Hamiltonian

$$H_P(J, \psi, p_\xi, \xi; \delta) = \frac{1}{2N} a(\delta) J^2 + \frac{p_\xi^2}{2} + \hat{U}_\delta(\psi, \xi). \quad (131)$$

Thus, as in Lemma 4.5, the Poincaré map of (121) is $o(\delta^{1/2})$ -close to the $O(\delta^{1/2})$ -time map for the Hamiltonian (131). Since $\hat{U}_\delta(\psi, \xi)$ depends continuously on δ with all derivatives, it follows from the non-degeneracy assumption Box2 and the KAM non-degeneracy assumption Box3 that the Hamiltonian (131) has a KAM-non-degenerate elliptic fixed point near $(J, \psi, p_\xi, \xi) = (0, 0, 0, \xi_{min})$ with a positive measure set of KAM tori around it. Hence the Poincaré map of the Hamiltonian (121) also has such elliptic fixed point, proving Theorem 4.

5.3 Smoothness properties of the averaged potential

Before proceeding to the proof of Theorem 5 (the case of simultaneous impacts) we investigate regularity properties of the averaged potential W_{avg} in (124). We start with the case $\delta = 0$, i.e., let the function q^* in (124) be the billiard's saw-tooth solution (34).

Lemma 5.1. *Provided $\|\xi^{(n)} - \xi^{(m)}\| > \rho$, the averaged potential along the saw-tooth solution, W_{avg} of (36), is C^∞ for $\theta^{(n)} - \theta^{(m)} \neq 0 \pmod{\pi}$. Under the parity assumption Box4, at the singularities $\theta^{(n)} - \theta^{(m)} = 0 \pmod{\pi}$, the potential W_{avg} is C^2 -smooth, yet it is not C^3 smooth in general.*

Proof. Setting the shorthand notation $\vartheta = \theta^{(n)} - \theta^{(m)}$ and $\zeta = \xi^{(n)} - \xi^{(m)}$, and using the periodicity of q^* , the averaged potential of (36) becomes

$$W_{avg}(\vartheta, \zeta) = \frac{1}{2\pi} \int_0^{2\pi} W(q^*(s) - q^*(s + \vartheta), \zeta) ds.$$

By (34), for $\vartheta \in (0, \pi)$,

$$\begin{aligned} 2\pi W_{avg}(\vartheta, \zeta) &= \int_0^{\pi-\vartheta} W(-\vartheta, \zeta) ds + \int_{\pi-\vartheta}^\pi W(2s - 2\pi + \vartheta, \zeta) ds \\ &\quad + \int_\pi^{2\pi-\vartheta} W(\vartheta, \zeta) ds + \int_{2\pi-\vartheta}^{2\pi} W(4\pi - 2s - \vartheta, \zeta) ds \\ &= (\pi - \vartheta)(W(-\vartheta, \zeta) + W(\vartheta, \zeta)) \\ &\quad + \int_{\pi-\vartheta}^\pi (W(2s - 2\pi + \vartheta, \zeta) + W(-2s + 2\pi - \vartheta, \zeta)) ds \end{aligned} \quad (132)$$

so for $\|\zeta\| > \rho$ it is C^∞ for $\vartheta \in (0, \pi)$. Now, recall that $W_{avg}(\vartheta, \zeta)$ is an even function, so its extension to $\vartheta \in (-\pi, 0)$ is simply $W_{avg}(\vartheta, \zeta) = W_{avg}(-\vartheta, -\zeta)$, hence it is C^∞ for all ϑ away from the matching points $\vartheta = 0 \pmod{\pi}$.

It remains to prove the C^2 -smoothness at the matching points. By the periodicity and the parity assumption Box4,

$$W_{avg}(\vartheta, \zeta) = W_{avg}(-\vartheta, \zeta) = W_{avg}(2\pi - \vartheta, \zeta),$$

so it is enough to verify only that the first derivatives of W_{avg} in (132) vanish at $\vartheta = 0, \pi$.

By the parity assumption the interaction potential W is even in ϑ , so (132) becomes

$$\begin{aligned} 2\pi W_{avg}(\vartheta, \zeta) &= 2(\pi - \vartheta)W(\vartheta, \zeta) + 2 \int_{\pi-\vartheta}^\pi W(2s - 2\pi + \vartheta, \zeta) ds \\ &= 2\pi W(\vartheta, \zeta) + 2 \int_0^\vartheta (W(-2u + \vartheta, \zeta) - W(\vartheta, \zeta)) du. \end{aligned} \quad (133)$$

Therefore,

$$\frac{\partial}{\partial \vartheta} W_{avg}(\vartheta, \zeta) = \frac{\partial}{\partial \vartheta} W(\vartheta, \zeta) + \frac{1}{\pi} \int_0^\vartheta \frac{\partial}{\partial \vartheta} (W(-2u + \vartheta, \zeta) - W(\vartheta, \zeta)) du = \left(1 - \frac{\vartheta}{\pi}\right) \frac{\partial}{\partial \vartheta} W(\vartheta, \zeta),$$

so $\frac{\partial}{\partial \vartheta} W_{avg}(\vartheta, \zeta)|_{\vartheta=0, \pi} = 0$, as required.

Differentiating further, we find

$$\frac{\partial^2}{\partial \vartheta^2} W_{avg}(\vartheta, \zeta) = \left(1 - \frac{\vartheta}{\pi}\right) \frac{\partial^2}{\partial \vartheta^2} W(\vartheta, \zeta) - \frac{1}{\pi} \frac{\partial}{\partial \vartheta} W(\vartheta, \zeta). \quad (134)$$

Next, we obtain

$$\frac{\partial^3}{\partial \vartheta^3} W_{avg}(\vartheta, \zeta) = \left(1 - \frac{\vartheta}{\pi}\right) \frac{\partial^3}{\partial \vartheta^3} W(\vartheta, \zeta) - \frac{2}{\pi} \frac{\partial^2}{\partial \vartheta^2} W(\vartheta, \zeta),$$

so $\frac{\partial^3}{\partial \vartheta^3} W_{avg}(\vartheta, \zeta)|_{\vartheta \rightarrow 0, \pi} = -\frac{2}{\pi} \frac{\partial^2}{\partial \vartheta^2} W(\vartheta, \zeta)|_{\vartheta \rightarrow 0, \pi}$. Generically, these values do not vanish, so the even extension of W_{avg} to negative ϑ cannot be C^3 (i.e., it is only piecewise smooth). \square

Let us now consider the case of $\delta > 0$. As in the lemma above, we use the notation $\vartheta = \theta^{(n)} - \theta^{(m)}$ and $\zeta = \xi^{(n)} - \xi^{(m)}$.

Lemma 5.2. *Away from $\vartheta = 0 \pmod{\pi}$, the averaged potential $W_{avg}(\vartheta, \zeta; \delta)$ is C^∞ -close, for small δ to the saw-tooth averaged potential $W_{avg}(\vartheta, \zeta; 0)$.*

When the parity assumption Box4 is satisfied, the averaged potential near $\vartheta = 0 \pmod{\pi}$ is C^2 -close to $W_{avg}(\vartheta, \zeta; 0)$, along with the derivatives with respect to ζ . The higher order derivatives with respect to ϑ do not, in general, have a continuous limit near the singular values $\vartheta = 0 \pmod{\pi}$ as $\delta \rightarrow 0$. The following estimates hold true:

$$\begin{aligned} \frac{\partial^3}{\partial \vartheta^3} W_{avg}(\vartheta, \zeta; \delta) &= O(1), \\ \frac{\partial^4}{\partial \vartheta^4} W_{avg}(\vartheta, \zeta; \delta) &= O(\delta^{-\frac{1}{\alpha}}), \\ \frac{\partial^5}{\partial \vartheta^5} W_{avg}(\vartheta, \zeta; \delta) &= O(\delta^{-\frac{2}{\alpha}}); \end{aligned}$$

differentiation with respect to ζ does not affect these estimates:

$$\frac{\partial^{i+k}}{\partial \vartheta^i \partial \zeta^k} W_{avg}(\vartheta, \zeta; \delta) = O(\delta^{-\frac{i-3}{\alpha}}), \quad i = 3, 4, 5; \quad k \geq 0.$$

Proof. In order to establish the regularity for $\vartheta \neq 0 \pmod{\pi}$ (i.e., for non-simultaneous impacts), we take $\vartheta \in (0, \pi)$ and let $\eta > 0$ be a sufficiently small number so that $\eta <$

$\min(\vartheta, \pi - \vartheta)$. By the periodicity of q_δ^* , we write (124) as

$$\begin{aligned}
2\pi W_{avg}(\vartheta, \zeta; \delta) &= \int_{-\eta}^{2\pi-\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds \\
&= (\int_{-\eta}^{\pi-\vartheta-\eta} + \int_{\pi-\vartheta+\eta}^{\pi} + \int_{\pi-\vartheta+\eta}^{2\pi-\vartheta-\eta} + \int_{2\pi-\vartheta-\eta}^{2\pi-\eta}) W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds \\
&= \int_{-\eta}^{\pi-\vartheta-\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds + \int_{\pi-\eta}^{\pi+\eta} W(q_\delta^*(s - \vartheta) - q_\delta^*(s), \zeta) ds \\
&\quad + \int_{\pi-\vartheta+\eta}^{2\pi-\vartheta-\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds + \int_{2\pi-\eta}^{2\pi+\vartheta-\eta} W(q_\delta^*(s - \vartheta) - q_\delta^*(s), \zeta) ds.
\end{aligned}$$

The function q_δ^* is C^∞ for all $\delta \geq 0$ when its argument is bounded away from $0 \pmod{\pi}$, so every integral in this sum is C^∞ function of ϑ for all $\delta \geq 0$ (this is an explicit version of a similar statement in Lemma 4.7). Similarly, one proves the regularity of W_{avg} for $\vartheta \in (\pi, 2\pi)$.

Let us now examine the case where ϑ is close to 0 or π . We have

$$2\pi W_{avg}(\vartheta, \zeta; \delta) = \int_0^{2\pi} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds := S(\vartheta, \zeta; \delta, \eta) + R(\vartheta, \zeta; \delta, \eta), \quad (135)$$

where S , the ‘‘singular part’’, corresponds to the integration intervals with both particles η -close to impacts:

$$S(\vartheta, \zeta; \delta, \eta) = \int_{-\eta}^{\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds + \int_{\pi-\eta}^{\pi+\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds, \quad (136)$$

and R , the ‘‘regular part’’, corresponds to the integration intervals for which both particles are at a distance larger than η from impacts:

$$R(\vartheta, \zeta; \delta, \eta) = \int_{\eta}^{\pi-\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds + \int_{\pi+\eta}^{2\pi-\eta} W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds. \quad (137)$$

For small η , and ϑ close to 0 (both particles have the same phase) or π (the anti-phase state), the term $R(\vartheta, \zeta; \delta, \eta)$ is C^∞ for all $\delta \geq 0$. Thus, we need to evaluate the singular term, S , and its derivatives. Recall that $q_\delta^*(s) \rightarrow q_0^*(s)$ in C^0 , so S converges in C^0 to the billiard limit $S(\vartheta, \zeta; 0, \eta)$. We will show that up to order 3 the derivatives of S are uniformly bounded for all small δ . This proves, by compactness argument, the C^2 -closeness of S and, hence, of W_{avg} , to their C^0 limits at $\delta = 0$, as claimed.

So, to prove the lemma, we only need to estimate the derivatives of S . In order to do this, we use formula (92) for q_δ^* , where the first impact is at $M^1 : q = 0$ (this corresponds to $\theta_1 = 0$) and the second impact is at $M^2 : q = \pi$ (this corresponds to $\theta_2 = \pi$). Recall that q_δ^* is the periodic orbit in the energy level $H = \frac{1}{2}$ of the system (118). Since this system has one degree of freedom and is reversible, the periodic orbit $q_\delta^*(\theta)$ is even and 2π -periodic. Moreover, by the symmetry of the potential $V_d(q)$ (see Assumption Box1), we have

$q_\delta^*(\theta) = q_\delta^*(\pi - \theta)$. Thus, we may write (92) as

$$\begin{aligned} q_\delta^*(\theta) &= \delta^{1/\alpha} \tilde{q}_\delta \left(\frac{\theta}{\omega_0(\delta) \delta^{1/\alpha}} \right) \quad \text{near } \theta = 0, \\ q_\delta^*(\theta) &= \pi - \delta^{1/\alpha} \tilde{q}_\delta \left(\frac{\theta - \pi}{\omega_0(\delta) \delta^{1/\alpha}} \right) \quad \text{near } \theta = \pi. \end{aligned} \quad (138)$$

where $\tilde{q}_\delta(\cdot)$ is an even function with bounded derivatives (uniformly for all $\delta \geq 0$). By (118)

$$\frac{d}{dt} q_\delta^* = p_\delta^*, \quad \frac{d}{dt} p_\delta^* = -\delta V_d'(q_\delta^*).$$

Therefore, if we denote

$$\tilde{p}_\delta \left(\frac{\theta}{\omega_0(\delta) \delta^{1/\alpha}} \right) = \tilde{q}'_\delta = \frac{d}{dt} q_\delta^* \quad (139)$$

(see (71) and (92)), then

$$\tilde{p}'_\delta = -\frac{\alpha \delta^{1+1/\alpha}}{Q_d(\delta^{1/\alpha} \tilde{q})^{\alpha+1}} Q_d'(\delta^{1/\alpha} \tilde{q}) = -\frac{1}{\tilde{q}_\delta^{\alpha+1}} \left(\frac{\alpha}{(Q_d'(0))^\alpha} + O(\delta^{1/\alpha}) \right). \quad (140)$$

Notice that by Lemma 4.1, the behavior of \tilde{q} is asymptotically linear at large u :

$$\tilde{q}_\delta(u) = |u| + o(u), \quad (141)$$

so

$$\tilde{p}_\delta^{(k)}(u) = O(|u|^{-\alpha-k}), \quad k \geq 1. \quad (142)$$

Note also that the frequency ω_0 of the vertical oscillations tends to that of the billiard motion, i.e.,

$$\omega_0(0) = 1. \quad (143)$$

Now we can return to analyze the behavior of the singular term S . For ϑ close to 0, using formulas (138) and that W is even, we write (136) as

$$\begin{aligned} S(\vartheta, \zeta; \delta, \eta) &= \left(\int_{-\eta}^{\eta} + \int_{\pi-\eta}^{\pi+\eta} \right) W(q_\delta^*(s) - q_\delta^*(s + \vartheta), \zeta) ds = 2 \int_{-\eta}^{\eta} W \left(\delta^{1/\alpha} \tilde{q}_\delta \left(\frac{s}{\omega_0 \delta^{1/\alpha}} \right) - \delta^{1/\alpha} \tilde{q}_\delta \left(\frac{s + \vartheta}{\omega_0 \delta^{1/\alpha}} \right), \zeta \right) ds \\ &= 2\omega_0 \delta^{1/\alpha} \int_{-\eta/(\omega_0 \delta^{1/\alpha})}^{\eta/(\omega_0 \delta^{1/\alpha})} W \left(\delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta \left(u + \frac{\vartheta}{\omega_0 \delta^{1/\alpha}} \right), \zeta \right) du. \end{aligned}$$

Similarly, near $\vartheta = \pi$, we have

$$\begin{aligned}
S(\vartheta, \zeta; \delta, \eta) &= \left(\int_{-\eta}^{\eta} + \int_{\pi-\eta}^{\pi+\eta} \right) W(q_{\delta}^*(s) - q_{\delta}^*(s + \vartheta), \zeta) ds \\
&= \int_{-\eta}^{\eta} W(\delta^{1/\alpha} \tilde{q}_{\delta}(\frac{s}{\omega_0 \delta^{1/\alpha}}) + \delta^{1/\alpha} \tilde{q}_{\delta}(\frac{s + \vartheta - \pi}{\omega_0 \delta^{1/\alpha}}) - \pi, \zeta) ds + \int_{\pi-\eta}^{\pi+\eta} W(\pi - \delta^{1/\alpha} \tilde{q}_{\delta}(\frac{s - \pi}{\omega_0 \delta^{1/\alpha}}) - \delta^{1/\alpha} \tilde{q}_{\delta}(\frac{s - 2\pi + \vartheta}{\omega_0 \delta^{1/\alpha}}), \zeta) ds \\
&= 2\omega_0 \delta^{1/\alpha} \int_{-\eta/(\omega_0 \delta^{1/\alpha})}^{\eta/(\omega_0 \delta^{1/\alpha})} W\left(\pi - \delta^{1/\alpha} \tilde{q}_{\delta}(u) - \delta^{1/\alpha} \tilde{q}_{\delta}(u + \frac{\vartheta - \pi}{\omega_0 \delta^{1/\alpha}}), \zeta\right) ds.
\end{aligned}$$

These two formulas can be written in a unified way:

$$S(\vartheta, \zeta; \delta, \eta) = 2\omega_0 \delta^{1/\alpha} \int_{-\eta/(\omega_0 \delta^{1/\alpha})}^{\eta/(\omega_0 \delta^{1/\alpha})} W\left(\sigma \pm \delta^{1/\alpha} \tilde{q}_{\delta}(u) - \delta^{1/\alpha} \tilde{q}_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta\right) du,$$

where one chooses $\sigma = 0$, $\hat{\vartheta} = \vartheta$ and the plus sign in front of $\delta^{1/\alpha} \tilde{q}_{\delta}(u)$ in the case of ϑ close to zero, and $\sigma = \pi$, $\hat{\vartheta} = \vartheta - \pi$ and the minus sign in front of $\delta^{1/\alpha} \tilde{q}_{\delta}(u)$ in the case of ϑ close to π .

The first derivative of S with respect to ϑ is

$$\frac{\partial}{\partial \vartheta} S(\vartheta, \zeta; \delta, \eta) = -2\delta^{1/\alpha} \int_{-\eta/(\omega_0 \delta^{1/\alpha})}^{\eta/(\omega_0 \delta^{1/\alpha})} \frac{\partial W}{\partial q}\left(\sigma \pm \delta^{1/\alpha} \tilde{q}_{\delta}(u) - \delta^{1/\alpha} \tilde{q}_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta\right) \tilde{p}_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du,$$

where \tilde{p}_{δ} is the derivative of \tilde{q}_{δ} (see (139)), so it is bounded with all derivatives by Lemma 4.1. Since the integrand is bounded (along with all derivatives with respect to ζ), it follows that $\frac{\partial S}{\partial \vartheta} = O(\eta)$, along with its derivatives with respect to ζ .

Next, we check the second derivative:

$$\begin{aligned}
\frac{\partial^2}{\partial \vartheta^2} S(\vartheta, \zeta; \delta, \eta) &= 2\omega_0^{-1} \delta^{1/\alpha} \int_{-\eta/(\omega_0 \delta^{1/\alpha})}^{\eta/(\omega_0 \delta^{1/\alpha})} \frac{\partial^2 W}{\partial q^2}(\sigma \pm \delta^{1/\alpha} \tilde{q}_{\delta}(u) - \delta^{1/\alpha} \tilde{q}_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^2 du \\
&\quad - 2\omega_0^{-1} \int_{-\eta/(\omega_0 \delta^{1/\alpha})}^{\eta/(\omega_0 \delta^{1/\alpha})} \frac{\partial W}{\partial q}(\sigma \pm \delta^{1/\alpha} \tilde{q}_{\delta}(u) - \delta^{1/\alpha} \tilde{q}_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}'_{\delta}(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du.
\end{aligned}$$

As before, the first line of the right-hand side is $O(\eta)$. Since \tilde{p}' decays as $|u|^{-\alpha-1}$ (see (142)), the integral in the second line is uniformly convergent. Thus, $\frac{\partial^2 S}{\partial \vartheta^2}$ is uniformly bounded for all small δ . The same is true for its derivatives with respect to ζ .

Differentiating further, we obtain that

$$\begin{aligned}
\frac{\partial^3}{\partial \vartheta^3} S(\vartheta, \zeta; \delta, \eta) &= -2\omega_0^{-2} \delta^{\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^3 W}{\partial q^3} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^3 du \\
&\quad + 6\omega_0^{-2} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad - 2\omega_0^{-2} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial W}{\partial q} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du.
\end{aligned} \tag{144}$$

As above, the first term in the right-hand side is $O(\eta)$. Integrating the last term by parts, we obtain

$$\begin{aligned}
\frac{\partial^3}{\partial \vartheta^3} S_j(\vartheta, \zeta; \delta, \eta) &= O(\eta) + \\
&\quad + 2\omega_0^{-2} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) (2\tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \pm \tilde{p}_\delta(u)) \tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad - 2\delta^{-\frac{1}{\alpha}} \left[\frac{\partial W}{\partial q} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \right]_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}
\end{aligned}$$

Since \tilde{p}'_δ decays as $|u|^{-\alpha-1}$, all the terms here are uniformly bounded, i.e., $\frac{\partial^3}{\partial \vartheta^3} S(\vartheta, \zeta; \delta, \eta)$ is uniformly bounded, with all its ζ -derivatives.

Similarly, differentiating (144) and integrating by parts, we obtain:

$$\begin{aligned}
\frac{\partial^4}{\partial \vartheta^4} S(\vartheta, \zeta; \delta, \eta) &= 2\omega_0^{-3} \delta^{\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^4 W}{\partial q^4} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^4 du \\
&\quad - 12\omega_0^{-3} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^3 W}{\partial q^3} (\delta^{1/\alpha} \sigma \pm \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^2 \tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad + 6\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) (\tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}))^2 du \\
&\quad + 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) (3\tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \pm \tilde{p}_\delta(u)) \tilde{p}''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad - 2\omega_0^{-3} \delta^{-\frac{2}{\alpha}} \left[\frac{\partial W}{\partial q} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \right]_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}
\end{aligned}$$

As before, the first term is $O(\eta)$. By (142), we have $\tilde{p}'_\delta(u) \approx 1/|u|^{\alpha+1}$ and $\tilde{p}''_\delta(u) \approx 1/|u|^{\alpha+2}$ at large $|u|$. This implies that the integrals in the 2-4 lines are uniformly bounded, so the second line is $O(1)$ and the third and fourth lines are $O(\delta^{-\frac{1}{\alpha}})$. The fifth line is of order $O(\delta^{-\frac{2}{\alpha}} \delta^{\frac{\alpha+2}{\alpha}}) = O(\delta)$. So the fourth order derivative with respect to ϑ diverges, along with its derivatives with respect to ζ , at most as $O(\delta^{-\frac{1}{\alpha}})$, in agreement with the claim of the lemma.

Namely

$$\begin{aligned}
\frac{\partial^4}{\partial \vartheta^4} S(\vartheta, \zeta; \delta, \eta) &= 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2} (\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \\
&\quad \times (3\tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}))^2 + (3\tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \pm \tilde{p}_\delta(u)) \tilde{p}''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du + O(1).
\end{aligned} \tag{145}$$

Finally, we evaluate the fifth derivative, using the same procedure as above (i.e., twice

differentiating (144), integrating by parts, and using the estimate (142) for the decay of the derivatives of \tilde{p} at large $|u|$):

$$\begin{aligned}
\frac{\partial^5}{\partial \vartheta^5} S(\vartheta, \zeta; \delta, \eta) &= -2\omega_0^{-4} \delta^{\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^5 W}{\partial q^5}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^5 du \\
&\quad + 20\omega_0^{-4} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^4 W}{\partial q^4}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^3 \tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad - 30\omega_0^{-4} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^3 W}{\partial q^3}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) (\tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}))^2 du \\
&\quad - 20\omega_0^{-4} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^3 W}{\partial q^3}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}})^2 \tilde{p}''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad + 20\omega_0^{-4} \delta^{-\frac{2}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}'_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \tilde{p}''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad + 2\omega_0^{-4} \delta^{-\frac{2}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) (4\tilde{p}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \pm \tilde{p}_\delta(u)) \tilde{p}'''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) du \\
&\quad - 2\omega_0^{-4} \delta^{-\frac{3}{\alpha}} \left[\frac{\partial W}{\partial q}(\sigma \pm \delta^{1/\alpha} \tilde{q}_\delta(u) - \delta^{1/\alpha} \tilde{q}_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}), \zeta) \tilde{p}'''_\delta(u + \frac{\hat{\vartheta}}{\omega_0 \delta^{1/\alpha}}) \right]_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \\
&= O(\delta^{-\frac{2}{\alpha}})
\end{aligned}$$

This completes the proof of the lemma. \square

Lemma 5.3. *Under the parity assumption, the second derivative of the averaged potential satisfies*

$$\frac{\partial^2}{\partial \vartheta^2} W_{avg}(0, \zeta; 0) = \frac{\partial^2 W}{\partial q^2}(0, \zeta), \quad \frac{\partial^2}{\partial \vartheta^2} W_{avg}(\pi, \zeta; 0) = -\frac{1}{\pi} \frac{\partial W}{\partial q}(\pi, \zeta). \quad (146)$$

The fourth derivative satisfies

$$\frac{\partial^4}{\partial \vartheta^4} W_{avg}(0, \zeta; \delta) = -\delta^{-\frac{1}{\alpha}} Q'_d(0) K(\alpha) \frac{\partial^2 W}{\partial q^2}(0, \zeta) (1 + o(1)_{\delta \rightarrow 0}), \quad (147)$$

$$\frac{\partial^4}{\partial \vartheta^4} W_{avg}(\pi, \zeta; \delta) = \delta^{-\frac{1}{\alpha}} Q'_d(0) K(\alpha) \frac{\partial^2 W}{\partial q^2}(\pi, \zeta) (1 + o(1)_{\delta \rightarrow 0}),$$

where

$$K(\alpha) = \int_{2^{1/\alpha}}^{\infty} \frac{4\alpha^2}{q^{2\alpha+2} \sqrt{1 - \frac{2}{q^\alpha}}} dq > 0. \quad (148)$$

Proof. Since $W_{avg}(\vartheta, \zeta; \delta)$ tends to the billiard limit in C^2 as $\delta \rightarrow 0$, formulas (146) are just given by (134). So, in order to prove (146), we only need to calculate the fourth order derivative. As shown in Lemma 5.2, the derivatives of the regular part (R in (135)) are uniformly bounded for all $\delta \geq 0$, and the fourth order derivative of W_{avg} at 0 and π is dominated by the derivative of S . It follows from (145) at $\vartheta = 0$ (hence, $\hat{\vartheta} = 0$ and $\sigma = 0$)

that

$$\begin{aligned} \frac{\partial^4}{\partial \vartheta^4} S(\vartheta, \zeta; \delta, \eta)|_{\vartheta=0} &= 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(0, \zeta) \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} (3\tilde{p}'_\delta(u)^2 + 4\tilde{p}_\delta(u)\tilde{p}''_\delta(u)) du + O(1) \\ &= -2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(0, \zeta) \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} (\tilde{p}'_\delta(u))^2 du + 8\omega_0^{-3} \frac{\partial^2 W}{\partial q^2}(0, \zeta) \delta^{-\frac{1}{\alpha}} [\tilde{p}_\delta(u)\tilde{p}'_\delta(u)]_{-\eta\omega_0 \delta^{-\frac{1}{\alpha}}}^{\eta\omega_0 \delta^{-\frac{1}{\alpha}}} + O(1). \end{aligned}$$

Since \tilde{p}'_δ decays sufficiently fast by (142),

$$\frac{\partial^4}{\partial \vartheta^4} S(\vartheta, \zeta; \delta, \eta)|_{\vartheta=0} = -2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(0, \zeta) \int_{-\infty}^{\infty} (\tilde{p}'_\delta(u))^2 du + O(1). \quad (149)$$

Using (140)

$$\int_{-\infty}^{\infty} (\tilde{p}'_\delta(u))^2 du = \frac{\alpha^2}{Q'_d(0)^{2\alpha}} \int_{-\infty}^{+\infty} \tilde{q}_0(u)^{-2\alpha-2} du + o(1), \quad (150)$$

where $(\tilde{q}_0(u), \tilde{p}_0(u))$ is the solution of the limit of the Hamiltonian system (118), namely

$$H = \frac{\tilde{p}^2}{2} + \frac{1}{(Q'_d(0)\tilde{q})^\alpha}$$

at $H = \frac{1}{2}$. Choosing symmetric parameterization of the time u (so $\tilde{p}_0(0) = 0$, hence $\tilde{q}_0(0) = 2^{1/\alpha}/Q'_d(0)$), we obtain

$$\int_{-\infty}^{+\infty} \tilde{q}_0(u)^{-2\alpha-2} du = 2 \int_{2^{1/\alpha}/Q'_d(0)}^{+\infty} \tilde{q}_0^{-2\alpha-2} \frac{d\tilde{q}_0}{\tilde{p}_0} = 2Q'_d(0)^{2\alpha+1} \int_{2^{1/\alpha}}^{+\infty} q^{-2\alpha-2} \sqrt{1 - \frac{2}{q^\alpha}} dq,$$

so

$$\int_{-\infty}^{\infty} (\tilde{p}'_\delta(u))^2 du = 2Q'_d(0)\alpha^2 \int_{2^{1/\alpha}}^{+\infty} q^{-2\alpha-2} \sqrt{1 - \frac{2}{q^\alpha}} dq + o(1).$$

Finally, by (149), and since $\omega_0(0) = 1$,

$$\frac{\partial^4}{\partial \vartheta^4} S(\vartheta, \zeta; \delta, \eta)|_{\vartheta=0} = -4\delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(0, \zeta) Q'_d(0)\alpha^2 \int_{2^{1/\alpha}}^{+\infty} q^{-2\alpha-2} \sqrt{1 - \frac{2}{q^\alpha}} dq + o(\delta^{-\frac{1}{\alpha}}),$$

and (147) follows at $\vartheta = 0$.

When $\vartheta = \pi$ (hence, $\hat{\vartheta} = 0$ and $\sigma = \pi$ in (145)), equation (145) gives

$$\begin{aligned} \frac{\partial^4}{\partial \vartheta^4} S(\pi, \zeta; \delta, \eta) &= 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2}(\pi - 2\delta^{1/\alpha} \tilde{q}_\delta(u), \zeta) (3(\tilde{p}'_\delta(u))^2 + 2\tilde{p}_\delta(u)\tilde{p}''_\delta(u)) du + O(1) \\ &= 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} \frac{\partial^2 W}{\partial q^2}(\pi, \zeta) (3(\tilde{p}'_\delta(u))^2 + 2\tilde{p}_\delta(u)\tilde{p}''_\delta(u)) du + O(1) \\ &= 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(\pi, \zeta) \int_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} (\tilde{p}'_\delta(u))^2 du + 4\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(\pi, \zeta) [\tilde{p}_\delta(u)\tilde{p}'_\delta(u)]_{-\eta/(\omega_0 \delta^{\frac{1}{\alpha}})}^{\eta/(\omega_0 \delta^{\frac{1}{\alpha}})} + O(1) \\ &= 2\omega_0^{-3} \delta^{-\frac{1}{\alpha}} \frac{\partial^2 W}{\partial q^2}(\pi, \zeta) \int_{-\infty}^{\infty} (\tilde{p}'_\delta(u))^2 du + O(1). \end{aligned}$$

The last line differs from (149) only by a factor, so (147) follows at $\vartheta = \pi$. \square

The above results allow us to characterize the behavior of the averaged potential $U(\theta, \xi; \delta)$ defined by (125). It is given by Lemma 5.2 that $U(\theta, \xi; \delta)$ depends on $\delta \geq 0$ continuously with the derivatives up to order 2. Therefore, by the non-degenerate minimum assumption Box2, it has a non-degenerate minimum line near the line (37) for all small δ :

$$\theta^{(n)} = \theta_{min}^{(n)}(\delta) + c, \quad \xi^{(n)} = \xi_{min}^{(n)}(\delta).$$

Lemma 5.4. *The potential $U(\theta, \xi; \delta)$ has the following expansion near the minimum line:*

$$\begin{aligned} U(\theta, \xi_{min}(\delta); \delta) &= \sum_{n=2}^N \sum_{m=1}^{n-1} \gamma_{nm}(\delta) (\theta^{(n)} - \theta^{(m)} - \theta_{min}^{(n)} - \theta_{min}^{(m)})^2 \\ &\quad + \delta^{-\frac{1}{\alpha}} K(\alpha) Q'_d(0) \beta_{nm}(\delta) (\theta^{(n)} - \theta^{(m)} - \theta_{min}^{(n)} - \theta_{min}^{(m)})^4 + \dots, \end{aligned} \quad (151)$$

where $\gamma_{nm}(\delta)$ and $\beta_{nm}(\delta)$ tend, as $\delta \rightarrow 0$, to γ_{nm} and, respectively, β_{nm} defined by (41); the positive coefficient $K(\alpha)$ is given by (148), and the dots stand for sixth and higher order terms of the expansion in powers of $(\theta^{(n)} - \theta^{(m)} - \theta_{min}^{(n)} - \theta_{min}^{(m)})$.

Proof. By (125), the derivatives of $U(\theta, \xi; \delta)$ with respect to $\theta^{(n)}$ are given by the sum over the corresponding derivatives of $W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}; \delta)$. The second derivatives tend to those of $W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}; 0)$ by Lemma 5.2, so the values of γ_{nm} follow from Lemma 5.3. Since $W_{avg}(\vartheta, \zeta; \delta)$ is even and 2π -periodic in ϑ , the third order derivatives with respect to ϑ vanish at $\vartheta = 0 \pmod{\pi}$, whereas the fourth order derivatives are given by Lemma 5.3, which determines the values of β_{nm} . \square

5.4 The simultaneous impacts case

Since the billiard choreographic solution in the case of simultaneous impacts satisfies, for any pair of particles, $\theta_{min}^{(n)} - \theta_{min}^{(m)} = 0 \pmod{\pi}$, the averaged potential in the limit $\delta = 0$ is not, in general C^3 smooth, as explained in Section 5.3. Therefore, the limit $\delta = 0$ is singular in this case. Indeed, we estimated the derivatives of the averaged potential up to order 5 and showed that starting with order 4 they tend to infinity as $\delta \rightarrow 0$.

To deal with this difficulty, we prove Theorem 5 by using different arguments from those used to prove Theorems 3 and 4. In Sections 5.4.1 and 5.4.2 we bring the system to the averaged form (122) and estimate the correction term $\delta\tilde{G}$. In Section 5.4.3 we bring the truncated averaged system to a Birkhoff normal form; the coefficients of the 4-th order terms in the normal form diverge as $\delta \rightarrow 0$, nevertheless we show that it has a KAM non-degenerate

elliptic periodic orbit for all small $\delta > 0$. In Section 5.4.4 we estimate the difference between the Birkhoff normal form for the full and truncated systems and, under the requirement $\alpha > 6$ show that the full system has an elliptic periodic orbit surrounded by KAM tori.

5.4.1 Expansion near the orbit $\mathbf{L}^*(\theta, \xi)$ of the uncoupled system.

Lemma 5.5. *In the case of simultaneous-impacts, provided $\alpha > 2$ and $\|\xi^{(n)} - \xi^{(m)}\| > \rho$ for all $n \neq m$, the expansion of the scaled Hamiltonian (121) near $\mathbf{L}^*(\theta, \xi)$ of (123) is of the form*

$$H = \sum_{n=1}^N \omega_0(\delta) I^{(n)} + \delta^{1/2} \sum_{n=1}^N \left(\frac{1}{2} a(\delta) (I^{(n)})^2 + \sum_{i=1}^{d-1} \left[\frac{(p_{\xi,i}^{(n)})^2}{2} + V_i(\xi_i^{(n)}) \right] \right) + \delta^{1/2} \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)}), \xi^{(n)} - \xi^{(m)}) + \delta \hat{G}(I, \theta, \xi; \delta), \quad (152)$$

where the derivatives of $\delta \hat{G}$ which include exactly k differentiations with respect to θ are bounded by $O(\delta^{1-k/\alpha})$.

Proof. In the box case, for $\alpha > 2$, the expansion (94) can be improved to

$$\hat{q}_d(\delta^{1/2} I, \theta; \delta) = q_\delta^*(\theta) + \delta^{1/2} G_d(I, \theta; \delta) \quad (153)$$

where derivatives of G_d which include exactly k differentiations with respect to θ are bounded by $O(\delta^{-k/\alpha})$. Indeed, since $I = 0$ corresponds to the periodic orbit L_δ^* , we have that $\hat{q}(0, \theta, 0) := q_\delta^*(\theta)$. Away from impacts $\hat{q}_d(\delta^{1/2} I, \theta; \delta)$ is a uniformly smooth function of $(\delta^{1/2} I, \theta)$, so (153) follows immediately. Near impacts, as the vertical dynamics are one-dimensional, equation (90) becomes, in the scaled coordinates (recall that here there are no \mathcal{P} variables):

$$\hat{q}_d(\delta^{1/2} I, \theta, \delta) = q_{impact} + \delta^{1/\alpha} \tilde{q} \left(\frac{\tau(\delta^{1/2} I, \theta; \delta) - t_{in}(\delta)}{\delta^{1/\alpha}}, \mathcal{E}(\delta^{1/2} I); \delta \right),$$

where q_{impact} is either 0 or π . Since (τ, \mathcal{E}) are smooth functions of (θ, I) , we can expand the above expression at $I = 0$:

$$\hat{q}_d(\delta^{1/2} I, \theta, \delta) = q_\delta^*(\theta) + \delta^{1/2} I \bar{G}_d(\theta, \delta^{1/2-1/\alpha} I; \delta).$$

Each derivative with respect to θ generates a $\delta^{-1/\alpha}$ factor whereas each derivative with respect to I generates a $\delta^{1/2-1/\alpha}$ factor. Hence, for $\alpha > 2$, the derivatives of $G_d = I \bar{G}_d$ which

include exactly k derivatives with respect to θ are bounded by $O(\delta^{-k/\alpha})$, as claimed.

Since $\|\xi^{(n)} - \xi^{(m)}\| > \rho$ for all $n \neq m$, the interaction potential W at any point along \mathbf{L}^* is smooth, so, plugging (153) in the interaction term of (121) gives

$$\begin{aligned} W(\hat{q}_d(\delta^{1/2}I^{(n)}, \theta^{(n)}; \delta) - \hat{q}_d(\delta^{1/2}I^{(m)}, \theta^{(m)}; \delta), \xi^{(n)} - \xi^{(m)}) &= W(q_\delta^*(\theta^{(n)}) - q_\delta^*(\theta^{(m)}), \xi^{(n)} - \xi^{(m)}) + \\ &+ \delta^{1/2} \hat{G}_d(\theta^{(n)}, \theta^{(m)}, \delta^{1/2}I^{(n)}, \delta^{1/2}I^{(m)}, \xi^{(n)} - \xi^{(m)}; \delta), \end{aligned}$$

where derivatives of \hat{G}_d which include exactly k derivatives with respect to $\theta^{(n)}$ or $\theta^{(m)}$ are of order $\delta^{-k/\alpha}$. Hence, (121), becomes of the required form (152), where \hat{G} is the sum of such functions \hat{G}_d and regular terms coming from the $O((\delta^{1/2}I^{(n)})^3)$ terms of (119). \square

5.4.2 Regularity of averaging

We average the Hamiltonian (152) by using the same transformation (111) as in the non-simultaneous impacts case. To this aim, we establish that Ψ of (111) and its derivatives have the correct regularity as $\delta \rightarrow 0$. We first show that

$$\Psi_1(u, u_0, v; \xi^{(n)} - \xi^{(m)}; \delta) = \int_{u_0}^u W(q_\delta^*(s+v) - q_\delta^*(s-v), \xi^{(n)} - \xi^{(m)}) ds \quad (154)$$

satisfies the following result (the weak regularity of Lemma 4.8 is replaced by weaker estimates on the derivatives, thus allowing both particles to visit simultaneously the impact regions):

Lemma 5.6. *Provided $\|\xi^{(n)} - \xi^{(m)}\| > \rho$, the derivatives of the function $\Psi_1(\frac{\theta^{(n)} + \theta^{(m)}}{2}, u_0, \frac{\theta^{(n)} - \theta^{(m)}}{2}, \xi^{(n)} - \xi^{(m)}; \delta)$ which include $k > 0$ differentiations with respect to $\theta^{(n)}, \theta^{(m)}$ are of order $\delta^{-(k-1)/\alpha}$.*

Proof. Since $\|\xi^{(n)} - \xi^{(m)}\| > \rho$, and the ξ variables are fixed along the integration interval, W is smooth and bounded function of its arguments. So all the derivatives with respect to ξ are bounded. When both particles are away from impacts (i.e. when the argument of q_δ^* is away from $\{0, \pi\}$), all the derivatives of q_δ^* are bounded. Otherwise, by (92), the k -th derivative of $q_\delta^*(\theta)$ is of order $\delta^{-(k-1)/\alpha}$, so, the same property is shared by $W(q_\delta^*(s + \frac{\theta^{(n)} - \theta^{(m)}}{2}) - q_\delta^*(s - \frac{\theta^{(n)} - \theta^{(m)}}{2}), \xi^{(n)} - \xi^{(m)})$ and its integral, Ψ_1 . \square

By (116), the function W_{avg} has the same regularity properties as Ψ_1 . Now perform the transformation defined by $\Psi(; \delta)$ of (111): $(I^{(n)}, p_\xi^{(n)}) \rightarrow (I^{(n)}, p_\xi^{(n)}) - \delta^{1/2}(\partial_{\theta^{(n)}}, \partial_{\xi^{(n)}})\Psi$. Notice that here (as opposed to the non-simultaneous impacts case) we apply the transformation to the non-truncated Hamiltonian (152) (i.e. we do not drop the correction term $\delta\hat{G}$). Though we formally obtain the same averaged system (122), now the $\delta\tilde{G}$ term includes the transformed $\delta\hat{G}$ term. The derivatives with respect to ξ and I of the $\delta\tilde{G}$ term remain of

order δ whereas its derivatives which include exactly k differentiations with respect to $\theta^{(n)}$, $n = 1, \dots, N$, are of order $\delta^{1-k/\alpha}$. In particular, for $\alpha > 6$ the correction term $\delta\tilde{G}$ in (122) and its derivatives up to order 3 are $o(\delta^{1/2})$.

5.4.3 Normal form of the truncated averaged system.

Let us study now the truncation of system (122), namely the system (128) for the case of simultaneous impacts. This system is translation-invariant (the Hamiltonian does not change when the same constant is added to all $\theta^{(n)}$). By Lemma 5.2, the potential $U_\delta(\theta, \xi)$ in (128) has, even for the simultaneous impacts case, a C^2 limit at $\delta = 0$. By the non-degeneracy assumption Box2, the potential U has a non-degenerate line of minima for all small δ . This line corresponds to an elliptic periodic orbit⁷ of system (128). Our goal here is to establish the KAM-nondegeneracy of it. This amounts to bringing the Poincaré map near the periodic orbit to the 4-th order Birkhoff normal form (Lemma 5.8 below) and verification of the twist condition (Lemma 5.9).

We introduce local normal coordinates $(\theta^{(n)}, I^{(n)}) \rightarrow (\varphi, \psi, P, J)$ near the periodic orbit in the same way as in (51),(52). Here $\varphi = \frac{1}{N}(\theta^{(1)} + \dots + \theta^{(N)})$, and P is symplectically conjugate to φ . The variables ψ vary near zero and are linear combinations of $((\theta^{(n)} - \theta_{min}^{(n)}) - (\theta^{(m)} - \theta_{min}^{(m)}))$, $n, m = 1, \dots, N$; the variables J are symplectically conjugate to ψ and also vary near zero. As in Section 5.2, the Poincaré return map of (128) at the level of fixed P , equals to the $O(\delta^{1/2})$ -map of the flow defined by the Hamiltonian

$$H_P(J, \psi, p_\xi, \xi; \delta) = \frac{1}{2N}a(\delta)J^2 + \frac{p_\xi^2}{2} + \hat{U}_\delta(\psi, \xi), \quad (155)$$

where \hat{U}_δ is the reduced potential of (129). This Hamiltonian is exactly of the same form as H_P of (131), yet, the regularity properties of the reduced potential $\hat{U}_\delta(\psi, \xi)$ are different in the simultaneous impacts case.

By the parity assumption Box4, $W_{avg}(\theta^{(n)} - \theta^{(m)}, \xi^{(n)} - \xi^{(m)}; \delta)$ is even and 2π -periodic in θ , so the phase differences along the minimum line remain locked at $\theta^{(n)} = 0 \pmod{\pi}$, $n = 1, \dots, N$, for all small δ . Therefore, as ψ is a linear function of $\theta^{(n)} - \theta^{(m)} - \theta_{min}^{(n)} - \theta_{min}^{(m)}$, the reduced potential is even in ψ :

$$\hat{U}_\delta(\psi, \xi) = \hat{U}_\delta(-\psi, \xi). \quad (156)$$

⁷This orbit is translation invariant, i.e., it is a relative equilibrium.

By this symmetry,

$$\frac{\partial^{j+l}}{(\partial\psi)^j(\partial\xi)^l}\hat{U}_\delta(\psi, \xi)|_{\psi=0, \xi=\xi_{min}} = 0 \quad \text{for odd } j. \quad (157)$$

In particular, $\frac{\partial^2}{\partial\psi\partial\xi}\hat{U}_\delta(\psi, \xi)|_{\psi=0, \xi=\xi_{min}} = 0$, so at $\psi = 0, \xi = \xi_{min}$ the quadratic part of the Hamiltonian (155) is block-diagonal. Therefore, the expansion of the Hamiltonian (155) up to order four terms is of the form

$$H_{4jet} = H_2^z(z, p; \delta) + H_2^\theta(\psi, J; \delta) + H_3^{\theta z}(\psi, z; \delta) + H_3^z(z; \delta) + H_4^z(z; \delta) + H_4^{\theta z}(\psi, z; \delta) + \delta^{-\frac{1}{\alpha}} H_4^\theta(\psi; \delta), \quad (158)$$

where, hereafter, $z = \xi - \xi_{min}, p = p_\xi$, and H_j denotes a homogeneous polynomial of order j . By Lemma 5.4, the coefficients of H_j are uniformly bounded for all $\delta \geq 0$.

Define the z -Hamiltonian:

$$H^z(z, p; \delta) = \frac{p^2}{2} + \hat{U}_\delta(0, \xi_{min} + z) = H_2^z(z, p; \delta) + H_3^z(z; \delta) + H_4^z(z; \delta) + O(z^5), \quad (159)$$

and the θ -Hamiltonian:

$$H^\theta(\psi, J; \delta) = \frac{1}{2N} a J^2 + \hat{U}_\delta(\psi, \xi_{min}) = H_2^\theta(\psi, J; \delta) + \delta^{-\frac{1}{\alpha}} H_4^\theta(\psi; \delta) + \dots, \quad (160)$$

where, hereafter, the dots stand for terms of order 5 and higher (with coefficients that may diverge as $\delta \rightarrow 0$). Then the Hamiltonian (155) is of the form

$$H_P = H^z(z, p; \delta) + H^\theta(\psi, J; \delta) + H_3^{\theta z}(\psi, z; \delta) + H_4^{\theta z}(\psi, z; \delta) + \dots, \quad (161)$$

where, by the symmetry $\psi \rightarrow -\psi$ (see (156)), the polynomials $H_{3,4}^{\theta z}(\psi, z; \delta)$ have only monomials which are quadratic in ψ .

Denote the fourth order Birkhoff normal form of the Hamiltonian (159) near the fixed point at the origin by $H_{NF}^z(z, p; \delta)$ and the Birkhoff normal form of the Hamiltonian (160) by $H_{NF}^\theta(\psi, J; \delta)$. Recall that in such normal forms all the terms up to order 4 are resonant – all non-resonant terms up to order 4 are eliminated by a sequence of symplectic coordinate transformations. Each of these transformations is a time-1 map of a certain polynomial Hamiltonian whose coefficients depend smoothly on the coefficients of the 4-jet of the original system, see e.g. Chapter 7 of [3]. It is a well-known fact that the normalizing transformations can be chosen such that the linear symmetries of the system are preserved. In particular, the symmetry $(\psi, J) \rightarrow -(\psi, J)$ survives the transformations we describe below.

Lemma 5.7. *There exists a symplectic transformation which depends continuously on $\delta > 0$*

and brings the Hamiltonian H^θ of (160) to the normal form

$$H_{NF}^\theta(\psi, J; \delta) = \tilde{H}_2^\theta(\psi, J; \delta) + \delta^{-\frac{1}{\alpha}} K(\alpha) \tilde{H}_4^\theta(\psi, J; \delta) + \dots, \quad (162)$$

where \tilde{H}_k^θ are homogeneous polynomials of (ψ, J) of degree k ($k = 2, 4$) with bounded coefficients depending continuously on δ . The polynomials $\tilde{H}_{2,4}^\theta$ have a limit at $\delta = 0$ such that the normal form of the Hamiltonian $H^{\theta,0}$ of Assumption Box5 is

$$H_{NF}^{\theta,0}(\psi, J) = \tilde{H}_2^\theta(\psi, J; 0) + \tilde{H}_4^\theta(\psi, J; 0) + \dots. \quad (163)$$

Proof. The normal form transformation we are going to describe is a composition of several symplectic transformations. First, we do a linear transformation which diagonalizes the quadratic part of the θ -Hamiltonian (160). By Lemma 5.4, the quadratic part of the Hamiltonian (160) is, in the limit $\delta = 0$, identical to the quadratic part of $H^{\theta,0}$. Therefore, by Assumption Box2, all the frequencies are distinct for small δ ; hence the quadratic part of (160) is indeed symplectically diagonalizable and the limit of the diagonalizing transformation at $\delta = 0$ also diagonalizes the quadratic part of $H^{\theta,0}$. Since the transformation is linear, it does not introduce third order terms.

Thus, the Hamiltonian (160) becomes $H^\theta(\psi, J; \delta) = \tilde{H}_2^\theta(\psi, J; \delta) + \delta^{-\frac{1}{\alpha}} K(\alpha) \hat{H}_4^\theta(\psi, J; \delta) + \dots$, where $\hat{H}_4^\theta(\psi, J; 0)$ coincides with the 4-th order terms of $H^{\theta,0}$ after the diagonalization. Next, we do a symplectic transformation in order to eliminate all non-resonant fourth order terms, i.e., to bring the θ -Hamiltonian to its normal form H_{NF}^θ . It is well-known (see [3], Chapter 7) that such normalizing transformation does not alter the resonant terms of order 4 or lower, hence, $\tilde{H}_4^\theta(\psi, J; \delta)$ is obtained from $\hat{H}_4^\theta(\psi, J; \delta)$ by throwing away the non-resonant terms. It follows that even though the normalizing transformation does not have a limit⁸ at $\delta = 0$, the term $\tilde{H}_4^\theta(\psi, J; \delta)$ has a limit, equal to the resonant part of $\hat{H}_4^\theta(\psi, J; 0)$, so the lemma follows. \square

Lemma 5.8. *The Birkhoff normal form of the reduced Hamiltonian (155) is the sum of the normal forms of the Hamiltonians (159) and (160) with bounded corrections of order 4 which vanish at $(J, \psi) = 0$:*

$$H_{NF} = H_{NF}^z(z, p; \delta) + H_{NF}^\theta(\psi, J; \delta) + H_2^{\psi J}(\psi, J; \delta) H_2^{zp}(z, p; \delta) + H_4^{\psi J}(\psi, J; \delta) + \dots. \quad (164)$$

⁸The transformation is the time-1 map, $(\psi, J) \rightarrow (\psi, J) + \delta^{-\frac{1}{\alpha}} K(\alpha) (\partial_J S, -\partial_\psi S) + O((\psi, J)^4)$, of a Hamiltonian flow defined by the Hamilton function $\delta^{-\frac{1}{\alpha}} K(\alpha) S(\psi, J; \delta)$, where S is a fourth-degree homogeneous polynomial with bounded coefficients which solves the equation $\{\tilde{H}_2^\theta, S\} = \tilde{H}_4^\theta - \hat{H}_4^\theta$. The normalizing transformation for $H^{\theta,0}$ is the time-1 map for the Hamiltonian $S(\psi, J; 0)$.

Proof. Consider the symplectic transformation which brings the Hamiltonian $H^z = H_2^z(z, p; \delta) + H_3^z(z; \delta) + H_4^z(z; \delta) + \dots$ to its Birkhoff normal form up to order 4:

$$H_{NF}^z(z, p; \delta) = \tilde{H}_2^z(z, p; \delta) + \tilde{H}_4^z(z, p; \delta) + \dots \quad (165)$$

Since, by Lemma 5.2, the coefficients of all monomials in H^z up to order 4 have well-defined finite limits as $\delta \rightarrow 0$, this transformation also has a well-defined finite limit. Applying this transformation to H_P of (155) (equivalently, to (161)), we obtain, in the new coordinates,

$$H_P = \tilde{H}_2^z(z, p; \delta) + \tilde{H}_4^z(z, p; \delta) + H_2^\theta(\psi, J; \delta) + \hat{H}_3^{\theta z}(\psi, z, p; \delta) + \hat{H}_4^{\theta z}(\psi, z, p; \delta) + \delta^{-1/\alpha} H_4^\theta(\psi; \delta) + \dots, \quad (166)$$

where all monomials in $\hat{H}_{3,4}^{\theta z}(\psi, z, p; \delta)$ are quadratic in ψ (because of the symmetry $\psi \rightarrow -\psi$). Since all the terms in $\hat{H}_3^{\theta z}(\psi, z, p; 0)$ are non-resonant by Assumption Box5 (as at the minimum of \hat{U} , the frequencies have no resonances of third order), they remain non-resonant for small δ . Therefore, there exists a symplectic transformation $(z, p) \rightarrow (z, p) + O((\psi, J)^2)$, $(\psi, J) \rightarrow (\psi, J) + O((z, p)) \cdot (\psi, J)$ which eliminates the cubic terms and brings the Hamiltonian H_P to the form

$$)H_P = \tilde{H}_2^z(z, p; \delta) + \tilde{H}_4^z(z, p; \delta) + H_2^\theta(\psi, J; \delta) + \bar{H}_4^{\theta z}(\psi, J, z, p; \delta) + \bar{H}_4^\theta(\psi, J; \delta) + \delta^{-1/\alpha} H_4^\theta(\psi; \delta) + \dots, \quad (167)$$

where $\bar{H}_4^{\theta z}(\psi, J, z, p; \delta)$ is a sum of monomials quadratic in $(\psi, J)^2$ and $(z, p)^2$.

Now, we apply to H_P the transformation of Lemma 5.7, which brings the Hamiltonian $H_2^\theta(\psi, J; \delta) + \delta^{-1/\alpha} H_4^\theta(\psi; \delta) + \dots$ to its Birkhoff normal form. This brings H_P to the form (164), with some of the 4th order terms in $H_2^{\psi J}$, $H_2^{z p}$, and $H_4^{\psi J}$ possibly non-resonant. All such non-resonant terms are eliminated by a symplectic transformation without changing all other terms up of order four or less. \square

Because there are no resonances up to order 4 by Assumption Box5, the 4-jets of the Birkhoff normal forms depend only on actions. Namely, we introduce actions

$$\begin{aligned} I_{z_j} &= \frac{1}{2}(z_j^2 + p_j^2), & j &= 1, \dots, N(d-1), \\ I_{\psi_k} &= \frac{1}{2}(\psi_k^2 + J_k^2), & k &= 1, \dots, N-1. \end{aligned} \quad (168)$$

Then

$$\begin{aligned} H_{NF}^z &= \omega_z(\delta) I_z + I_z^T B_z(\delta) I_z + \dots, \\ H_{NF}^\theta &= \omega_\psi(\delta) I_\psi + \delta^{-1/\alpha} K(\alpha) I_\psi^T B_\psi(\delta) I_\psi + \dots, \end{aligned}$$

where, by Lemmas 5.2 and 5.7 $\omega_z(\delta)$, $\omega_\psi(\delta)$, $B_z(\delta)$, $B_\psi(\delta)$ all are continuous functions with

bounded limits at $\delta = 0$. By Assumption Box5, the determinants of $B_z(0)$ and $B_\psi(0)$ are non-zero.

Moreover, by Lemma 5.7, the Hamiltonian $H_{NF}^{\theta,0}$ is of the form

$$H_{NF}^{\theta,0} = \omega_\psi(0)I_\psi + I_\psi^T B_\psi(0)I_\psi + \dots .$$

Also, by Lemma 5.8, the Birkhoff normal form H_{NF} of (155) is given by

$$H_{NF} = H_{NF}^z + H_{NF}^\psi + O(\|I_z\| + \|I_\psi\|) \cdot I_\psi + \dots .$$

Lemma 5.9. *The Birkhoff normal form of the Hamiltonian (155) is non-degenerate.*

Proof. We need to check that the twist condition for H_{NF} is satisfied. Indeed,

$$\begin{aligned} \det\left(\frac{\partial^2 H_{NF}}{\partial I^2}\right) &= \begin{vmatrix} B_z(\delta) & O(1) \\ O(1) & \delta^{-1/\alpha} K(\alpha) B_\psi(\delta) + O(1) \end{vmatrix} = \delta^{-2N/\alpha} K(\alpha)^{2N} \begin{vmatrix} B_z(\delta) & O(1) \\ O(\delta^{1/\alpha}) & B_\psi(\delta) + O(\delta^{1/\alpha}) \end{vmatrix} \\ &= \delta^{-2N/\alpha} K(\alpha)^{2N} (|B_z(\delta)| \cdot |B_\psi(\delta)| + O(\delta^{1/\alpha})). \end{aligned} \tag{169}$$

Since the determinants $|B_z(\delta)|$ and $|B_\psi(\delta)|$ are bounded away from zero for all sufficiently small δ , we have $\det(\frac{\partial^2 H_{NF}}{\partial I^2}) \neq 0$, i.e. the twist condition is satisfied, so H_{NF} is non-degenerate. \square

5.4.4 Normal form of the full system.

We have shown that the truncation of system (122), namely system (128) for the case of simultaneous impacts, has a stable KAM non-degenerate elliptic periodic orbit for all $\delta > 0$. To complete the proof of Theorem 5, we need to show that the periodic orbit persists in the full system (122), which includes the error term $\delta\tilde{G}$, and also remains KAM-non-degenerate. This is non-trivial, as both the normal form and the error term $\delta\tilde{G}$ diverge in C^4 as $\delta \rightarrow 0$. While the persistence of the periodic orbit follows from the C^2 regularity of the normal form and the error term, the KAM non-degeneracy is established by revisiting the normal form computation of Section 5.4.3 applied to the full system.

Rewriting (122) with the use of the coordinates $(P, \varphi, J, \psi, p_\xi, \xi)$ of (130), we obtain the perturbed Hamiltonian:

$$\begin{aligned} H(P, \varphi, J, \psi, p_\xi, \xi; \delta) &= \omega_0 P + \delta^{1/2} \left(\frac{1}{2N} a P^2 + \frac{1}{2N} a J^2 + \frac{p_\xi^2}{2} + \hat{U}_\delta(\psi, \xi) \right) \\ &\quad + \delta\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta). \end{aligned} \tag{170}$$

The corresponding system is:

$$\begin{aligned}
\dot{\varphi} &= \omega_0 + \frac{1}{N}a\delta^{1/2}P + \delta\frac{\partial}{\partial P}\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta), \\
\dot{P} &= -\delta\frac{\partial}{\partial \varphi}\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta), \\
\dot{\psi} &= \frac{1}{N}a\delta^{1/2}J + \delta\frac{\partial}{\partial J}\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta), \\
\dot{J} &= -\delta^{1/2}\frac{\partial}{\partial \psi}\hat{U}_\delta(\psi, \xi) - \delta\frac{\partial}{\partial \psi}\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta), \\
\dot{\xi} &= \delta^{1/2}p_\xi + \delta\frac{\partial}{\partial p_\xi}\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta), \\
\dot{p}_\xi &= -\delta^{1/2}\frac{\partial}{\partial \xi}\hat{U}_\delta(\psi, \xi) - \delta\frac{\partial}{\partial \xi}\tilde{G}(P, \varphi, J, \psi, p_\xi, \xi; \delta).
\end{aligned} \tag{171}$$

Restricting to a given energy level (so P is determined by all other variables) and using $\bar{\varphi} = \delta^{1/2}\varphi$ as a new time:

$$\begin{aligned}
\frac{d\psi}{d\bar{\varphi}} &= \frac{1}{N\omega_0}aJ + \delta^{1/2}G_1(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta), \\
\frac{dJ}{d\bar{\varphi}} &= -\frac{1}{\omega_0}\frac{\partial}{\partial \psi}\hat{U}_\delta(\psi, \xi) - \delta^{\frac{1}{2}-\frac{1}{\alpha}}G_2(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta), \\
\frac{d\xi}{d\bar{\varphi}} &= \frac{1}{\omega_0}p_\xi + \delta^{1/2}G_3(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta), \\
\frac{dp_\xi}{d\bar{\varphi}} &= -\frac{1}{\omega_0}\frac{\partial}{\partial \xi}\hat{U}_\delta(\psi, \xi) - \delta^{1/2}G_4(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta),
\end{aligned} \tag{172}$$

where $\frac{\partial}{\partial \psi^k}G_i(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta) = O(\delta^{-k/\alpha})$, see the end of Section 5.4.2. The reduced system (172) is a fast-oscillating non-autonomous Hamiltonian system with the Hamilton function

$$\begin{aligned}
H_{reduced}(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta) &= \frac{1}{2N\omega_0}aJ^2 + \frac{p_\xi^2}{2\omega_0} + \frac{1}{\omega_0}\hat{U}_\delta(\psi, \xi) + \delta^{1/2}G_5(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta) \\
&= \frac{1}{\omega_0}H_P(J, \psi, p_\xi, \xi; \delta) + \delta^{1/2}G_5(\bar{\varphi}, J, \psi, p_\xi, \xi; \delta),
\end{aligned} \tag{173}$$

where H_P is the Hamiltonian (155). As above, G_5 is a function with bounded derivatives with respect to J, p_ξ, ξ , whereas making k differentiations of G_5 with respect to ψ introduces a factor of $O(\delta^{-k/\alpha})$. Now, as apposed to (155), $H_{reduced}$ is a periodic function of the time variable $\bar{\varphi}$, with the period $2\pi\delta^{1/2}$.

Lemma 5.10. *Provided that $\alpha > 6$, the system (172) has a KAM-nondegenerate elliptic periodic orbit.*

Proof. As we mentioned in Section 5.4.3, due to the non-resonance assumption of Box2 and the continuous dependence of the quadratic part of the Hamiltonian $H_P(J, \psi, p_\xi, \xi; \delta)$ on δ , the truncated system (155) has a non-resonant elliptic fixed point at the origin. Since for $\alpha > 4$ the perturbation to (155), i.e. the term $\delta^{1/2}G_5$ in the Hamiltonian (173), is C^2 -small, the system (172) has, for all small δ , an elliptic periodic orbit $(J, \psi, p_\xi, \xi) = \delta^{\frac{1}{2}-\frac{1}{\alpha}}X_p(\bar{\varphi}; \delta)$ where $X_p(\bar{\varphi}; \delta)$ is a continuous bounded function for all $\delta \geq 0$. Let $(J, \psi, p_\xi, \xi) = \delta^{\frac{1}{2}-\frac{1}{\alpha}}X_p(\bar{\varphi}; \delta) + X$.

Then

$$H(\bar{\varphi}, X; \delta) := H_{reduced}(\delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta) + X; \delta) = H_P(X; \delta) + G_6(\bar{\varphi}, X; \delta), \quad (174)$$

where the non-autonomous term G_6 is given by

$$G_6(\bar{\varphi}, X; \delta) = H_P(\delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta) + X; \delta) - H_P(X; \delta) + \delta^{\frac{1}{2}} G_5(\bar{\varphi}, \delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta) + X; \delta).$$

Next we show that if $\alpha > 6$, the term $G_6(\bar{\varphi}, X; \delta)$ and its derivatives with respect to X including up to 3 derivatives with respect to ψ are small, while its 4th derivatives with respect to ψ are $o(\delta^{-1/\alpha})$:

$$\frac{\partial^{k+l}}{\partial \psi^k \partial (J, p_\xi, \xi)^l} G_6 = O(\delta^{\frac{1}{2}-\frac{k}{\alpha}}), \quad k = 0, 1, 2, 3, 4. \quad (175)$$

Indeed, first recall that $\frac{\partial}{\partial \psi^k} G_5$ is $O(\delta^{-\frac{k}{\alpha}})$, so for $\alpha > 6$ the term $\delta^{\frac{1}{2}} G_5$ gives a correct contribution to the estimates (175). Next, by Lagrange formula,

$$H_P(\delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta) + X; \delta) - H_P(X; \delta) = \delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta) \cdot \int_0^1 \nabla H_P(X + s \delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta); \delta) ds.$$

Recall that by Lemma 5.2 the potential $\hat{U}_\delta(\psi, \xi)$, hence the Hamiltonian H_P , has bounded derivatives up to order 3, yet its fourth derivatives with respect to ψ are of order $\delta^{-\frac{1}{\alpha}}$ and its fifth derivatives are of order $\delta^{-\frac{2}{\alpha}}$. Hence, by (176), $\frac{\partial}{\partial \psi^k} (H_P(\delta^{\frac{1}{2}-\frac{1}{\alpha}} X_p(\bar{\varphi}; \delta) + X; \delta) - H_P(X; \delta))$ for $k = 0, 1, 2$ is of order $O(\delta^{\frac{1}{2}-\frac{1}{\alpha}})$, whereas for $k = 3$ the derivatives are of order $\delta^{\frac{1}{2}-\frac{2}{\alpha}}$, and for $k = 4$ they are of order $\delta^{\frac{1}{2}-\frac{3}{\alpha}}$, all in agreement with (175).

In order to align with the notations of Section 5.4.3, we denote hereafter the coordinates of X by (J, ψ, p, z) . Then, by (161) and (174)

$$\begin{aligned} H(\bar{\varphi}, X; \delta) &= H_2^z(z, p; \delta) + H_3^z(z; \delta) + H_4^z(z; \delta) + H_2^\theta(\psi, J; \delta) + H_3^{\theta z}(\psi, z; \delta) \\ &\quad + H_4^{\theta z}(\psi, z; \delta) + \delta^{-1/\alpha} H_4^\theta(\psi; \delta) + G_6(\bar{\varphi}, X; \delta), \end{aligned} \quad (176)$$

where the functions H_j of (161) are homogeneous polynomials of degree j with bounded coefficients (continuous in δ) and $H_{3,4}^{\theta z}(\psi, z; \delta)$ have only monomials quadratic in ψ .

Recall that H_P was brought to its normal form by a sequence of coordinate transformations described in Lemma 5.8. Next, we apply similar transformation to the full Hamiltonian $H(\bar{\varphi}, X; \delta)$ of (174). The difference is that H is periodic in time $\bar{\varphi}$, so the normalizing transformations are also periodic in $\bar{\varphi}$. It is well-known that one can make the coordinate transformations such that the resulting Hamiltonian will become autonomous up to any

given order in X , i.e. such a transformation corresponds to averaging to any given order [3]. Our goal is to bring $H(\bar{\varphi}, X; \delta)$ to an autonomous normal form up to order 4 and compare it with the 4-th order normal form H_{NF} (see (164)) of the autonomous Hamiltonian H_P . Since the period is small, of order $\delta^{1/2}$, there are no additional resonant terms due to the dependence on $\bar{\varphi}$. Still, one needs to check that the singular terms in (176) do not destroy the twist condition.

Let us make the linear symplectic transformations which bring the Hamiltonians $H_2^\theta(\psi, J; \delta)$ and $H_2^z(z, p; \delta)$ to the diagonal form $\tilde{H}_2^\theta(\psi, J; \delta)$ and $\tilde{H}_2^z(z, p; \delta)$ as in (162) and (165), respectively. Applying these transformations to (176), we obtain a new Hamiltonian

$$\begin{aligned} H(\bar{\varphi}, X; \delta) = & \tilde{H}_2^z(z, p; \delta) + H_3^z(z, p; \delta) + H_4^z(z, p; \delta) + \tilde{H}_2^\theta(\psi, J; \delta) + H_3^{\theta z}(\psi, J, z, p; \delta) \\ & + H_4^{\theta z}(\psi, J, z, p; \delta) + \delta^{-1/\alpha} H_4^\theta(\psi, J; \delta) + G_7(\bar{\varphi}, X; \delta), \end{aligned} \quad (177)$$

where H_j are different functions from those in (176), yet they keep the same structure - they are homogeneous polynomials of degree j with bounded coefficients (continuous in δ), and $H_{3,4}^{\theta z}$ have only monomials quadratic in (ψ, J) . Note also that since the linear normalizing transformations have bounded coefficients for all $\delta \geq 0$, the derivatives of the function G_7 are of the same order as the derivatives of G_6 , as given by (175).

Since $X = 0$ is a periodic orbit of the Hamiltonian $H(\bar{\varphi}, X; \delta)$, it follows that the expansion of $G_7(\bar{\varphi}, X; \delta)$ in powers of X starts with quadratic terms. Moreover, the terms of the expansion which are independent of (ψ, J) are of order $\delta^{1/2}$, the terms linear in (ψ, J) are of order $\delta^{1/2-1/\alpha}$, and so on, e.g. the 4-th order terms in (ψ, J) are $O(\delta^{1/2-4/\alpha})$, i.e., $o(\delta^{-1/\alpha})$ since $\alpha > 6$. Therefore, expanding G_7 in powers of X we can rewrite (177) as

$$\begin{aligned} H(\bar{\varphi}, X; \delta) = & \tilde{H}_2^z(z, p; \delta) + \tilde{H}_2^\theta(\psi, J; \delta) + \hat{H}_2^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) \\ & + \hat{H}_3^z(\bar{\varphi}, z, p; \delta) + \hat{H}_3^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) + \hat{H}_4^z(\bar{\varphi}, z, p; \delta) + \hat{H}_4^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) \\ & + \delta^{-1/\alpha} \hat{H}_4^\theta(\bar{\varphi}, \psi, J; \delta) + \dots, \end{aligned} \quad (178)$$

where, hereafter, the dots stand for the terms of order higher than 4 (i.e., they are irrelevant for our purposes). The terms \hat{H}_j are homogeneous polynomials of X of degree j with periodic in $\bar{\varphi}$ coefficients, continuous and bounded for all $\delta \geq 0$. Moreover, $\hat{H}_2^{\theta z}$ is $O(\delta^{\frac{1}{2}-\frac{2}{\alpha}})$ -close to zero. The polynomials \hat{H}_3^z and \hat{H}_4^z are the cubic and, respectively quartic part of $H(\bar{\varphi}, X; \delta)$ at $(\psi, J) = 0$, so they are $O(\delta^{1/2})$ -close to $H_{3,4}^z(z, p; \delta)$ of (177). The polynomial $\hat{H}_3^{\theta z}$ is $O(\delta^{\frac{1}{2}-\frac{3}{\alpha}})$ -close to $H_3^{\theta z}$ of (177) (indeed, it has a linear in (ψ, J) part which is $O(\delta^{\frac{1}{2}-\frac{1}{\alpha}})$ -close to zero, since $H_3^{\theta z}$ has no linear part in ψ , a quadratic in (ψ, J) part which is $O(\delta^{\frac{1}{2}-\frac{2}{\alpha}})$ -close to $H_3^{\theta z}$, and the cubic in (ψ, J) part is, again, $O(\delta^{\frac{1}{2}-\frac{3}{\alpha}})$ -close to zero). Similarly, the polynomials $\hat{H}_4^{\theta z}$ and \hat{H}_4^θ are $O(\delta^{\frac{1}{2}-\frac{3}{\alpha}})$ -close to $H_4^{\theta z}$ and H_4^θ , respectively. Since $\alpha > 6$, all

these corrections are small, i.e., they vanish at $\delta = 0$.

Since $\hat{H}_2^{\theta z} = O(\delta^{1/2-2/\alpha})$, the $\bar{\varphi}$ -periodic, linear symplectic transformation of the variables X which diagonalizes the quadratic part ($\tilde{H}_2^z + \tilde{H}_2^\theta(\psi, J; \delta) + \hat{H}_2^{\theta z}$) of H and makes the quadratic part independent of $\bar{\varphi}$ is $O(\delta^{\frac{1}{2}-\frac{2}{\alpha}})$ -close to identity: $X \rightarrow (Id + \delta^{\frac{1}{2}-\frac{2}{\alpha}} M(\bar{\varphi}; \delta))X$, for some periodic in $\bar{\varphi}$ matrix M , continuous and bounded for all $\delta \geq 0$.

After this transformation, the Hamiltonian (178) becomes

$$\begin{aligned} H(\bar{\varphi}, X; \delta) = & \hat{H}_2^z(z, p; \delta) + \hat{H}_2^\theta(\psi, J; \delta) + \hat{H}_3^z(\bar{\varphi}, z, p; \delta) + \hat{H}_3^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) + \hat{H}_4^z(\bar{\varphi}, z, p; \delta) \\ & + \hat{H}_4^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) + \delta^{-1/\alpha} \hat{H}_4^\theta(\bar{\varphi}, \psi, J; \delta) + \dots, \end{aligned} \quad (179)$$

where \hat{H}_2^z and \hat{H}_2^θ are $O(\delta^{\frac{1}{2}-\frac{2}{\alpha}})$ -close to \tilde{H}_2^z and, respectively, \tilde{H}_2^θ ; the new homogeneous third degree polynomials \hat{H}_3^z and $\hat{H}_3^{\theta z}$ acquire $O(\delta^{\frac{1}{2}-\frac{2}{\alpha}})$ -corrections in comparison with \hat{H}_3^z and $\hat{H}_3^{\theta z}$ of (178), and the new homogeneous fourth degree polynomials \hat{H}_4^z , $\hat{H}_4^{\theta z}$, and \hat{H}_4^θ acquire $O(\delta^{\frac{1}{2}-\frac{3}{\alpha}})$ -corrections. It is only important for us that these corrections vanish at $\delta = 0$.

Next, we follow the same steps as in Lemma 5.8. We make a periodic in $\bar{\varphi}$, symplectic transformation of the (z, p) -coordinates which brings the Hamiltonian $\hat{H}_2^z(z, p; \delta) + \hat{H}_3^z(\bar{\varphi}, z, p; \delta) + \hat{H}_4^z(\bar{\varphi}, z, p; \delta)$ to its Birkhoff normal form, independent of $\bar{\varphi}$ up to order 4 (recall that no new resonances can be created here by the $\bar{\varphi}$ -dependence because the frequency is large). This Hamiltonian is close to the z -Hamiltonian of (159), therefore the resulting normal form $\hat{H}_{NF}^z(z, p; \delta)$ is close to the normal form H_{NF}^z , see (165); in particular, $\hat{H}_{NF}^z = \hat{H}_2(z, p; \delta) + \hat{H}_4(z, p; \delta)$ does not contain cubic terms, its quadratic and quartic terms $\hat{H}_{2,4}^z$ depend only on the actions I_z (see (168)), and their limit as $\delta \rightarrow 0$ coincides with the limit of $\tilde{H}_{2,4}^z$.

After this transformation Hamiltonian (177) becomes

$$\begin{aligned} H(\bar{\varphi}, X; \delta) = & \hat{H}_2^z(z, p; \delta) + \hat{H}_4^z(z, p; \delta) + \hat{H}_2^\theta(\psi, J; \delta) + \hat{H}_3^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) \\ & + \hat{H}_4^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) + \delta^{-1/\alpha} \hat{H}_4^\theta(\bar{\varphi}, \psi, J; \delta) + \dots, \end{aligned} \quad (180)$$

where the modified functions $\hat{H}_{3,4}$ have the same structure as before and are close to their counterparts in (166). Since the third order terms in the Hamiltonian are all non-resonant (by Assumption Box 5 and by the fact that the period in $\bar{\varphi}$ is small), they are eliminated by a normalizing symplectic transformation of X , which equals to the identity plus higher order $\bar{\varphi}$ -dependent terms and is close to that employed in Lemma 5.8 (the transformation

from (166) to (167)). The Hamiltonian (180) becomes

$$\begin{aligned}
H(\bar{\varphi}, X; \delta) &= \hat{H}_2^z(z, p; \delta) + \hat{H}_4^z(\varphi, z, p; \delta) + \hat{H}_2^\theta(\psi, J; \delta) \\
&\quad + \hat{H}_4^{\theta z}(\bar{\varphi}, \psi, J, z, p; \delta) + \delta^{-1/\alpha} \hat{H}_4^\theta(\bar{\varphi}, \psi, J; \delta) + \dots,
\end{aligned} \tag{181}$$

where \hat{H}_4 , the fourth-degree polynomials in X , are close to their counterparts in (167) (note that applying the transformation to the singular term, $\delta^{-1/\alpha} \hat{H}_4^\theta$ of (179), introduces additional singularities but only to the terms of order 5 or higher in (181)).

The last step is to bring the 4-th order terms to the autonomous normal form. This is done by a symplectic transformation and the result is equivalent to throwing away the non-resonant 4-th order terms and taking the average of the resonant ones over $\bar{\varphi}$. Since all terms \hat{H} in (181) coincide with their counterparts in (167) at $\delta = 0$, we immediately obtain that the resulting 4-th order Birkhoff normal form of the Hamiltonian (173) is

$$H_{NF4} = H_2^z(z, p; \delta) + H_4^z(z, p; \delta) + H_2^\theta(\psi, J; \delta) + \delta^{-1/\alpha} K(\alpha) H_4^\theta(\psi, J; \delta) + H_4^{\theta z}(\psi, J, z, p; \delta), \tag{182}$$

where $H_j^x(\cdot; \delta)$ have bounded coefficients, and, as $\delta \rightarrow 0$ they approach the corresponding terms of (164). Hence, the KAM-nondegeneracy of the elliptic orbit of the system (172) at $X = 0$ follows from Lemma 5.9. \square

This completes the proof of Theorem 5.

The divergence as $\delta \rightarrow 0$ of the terms of order 5 and higher in X does not alter the KAM nondegeneracy result, yet, it implies that quantitative estimates regarding the size of the stability island require analysis of the asymptotic δ -dependence of such terms.

6 Discussion

While the lack of ergodicity in Hamiltonian systems is expected, here we found a specific mechanism for breaking the ergodicity, which persists for arbitrarily high energy for any finite number of particles that interact by repelling forces. We constructed coherent states of the multi-particle gas that correspond to collision-free choreographic solutions which are stabilized at high energies. We have also built similar solutions for systems of weakly interacting particles and systems of attracting particles.

Let us list several future research directions.

Non-ergodicity of the gas of repelling particles in containers of dispersive geometry. We have established the existence of KAM-stable choreographic solutions for generic multi-particle systems in containers that support a stable periodic billiard motion

of one particle. Conjecturally, this includes any generic container with a sufficiently smooth convex boundary. Yet, there are open classes of billiards with piece-wise smooth boundaries which do not allow for stable single-particle motions – the main example is given by dispersive billiards whose boundary is built of strictly concave smooth pieces. We propose that it should be possible to apply our method for finding KAM-stable choreographies for such containers as well. Indeed, smoothing the billiard potential destroys the hyperbolic structure of the dispersive billiard, and there are several known mechanisms for creating elliptic, KAM non-degenerate periodic motions of a single particle in the billiard-like Hamiltonian (21) at arbitrary small δ for the case of dispersive $d(\geq 2)$ -dimensional containers [51, 52, 36, 38]. It is conjectured that such islands appear for dispersive billiard-like Hamiltonians generically [51]. The phase space volume of the islands in these cases vanishes with δ , yet, its scaling with δ is known, and depends on the asymptotic behavior of the container smoothing potential, V , near the boundary. Our techniques imply that such stability islands of the single particle motion can produce also choreographic solutions of the N particle system provided the perturbations induced by the interaction potential, W , are much smaller than the islands size. Hence, the results regarding the existence of KAM-stable choreographic solutions in containers that support stable motion can probably be extended to any container.

Physical relevance of the coherent states. Under which conditions can the constructed choreographic solutions be observed in realistic multi-particle systems? General estimates on the probability for an initial condition to belong to a KAM-torus of a Hamiltonian system with N -degrees of freedom are quite pessimistic even for small N [12, 13], but examples of such systems where stability islands are well-noticeable are also known [2, 45], e.g. in numerical experiments of [38] the islands are seen for N as large as 20. Therefore, Nekhoroshev-type estimates on the life time of the coherent states are, probably, most relevant for the physical realizability question. Sufficiently long living (i.e., effectively stable) coherent states can be of direct interest when N is not very large. For example, for a gas in a three-dimensional rectangular box, the construction of $N < (l_1 l_2)/\rho^2$ particles moving vertically in synchrony, as in Theorem 5, corresponds to pulsating fronts. For a gas in a convex container, the states of $N < |L^*|/\rho$ particles constructed in Theorem 3 corresponds to rings of current of a specific non-trivial spatial form. For each such state, the effective stability imposes limitations on N and the parameters of the system. Additional classes of choreographic solutions may allow to study coherent states with a larger number of particles.

Multi-path choreographic solutions provide such a class. When the single-particle system is non-integrable and has an elliptic periodic orbit, it has, typically, infinitely many elliptic orbits (e.g. around an elliptic periodic orbit there are typically many resonant elliptic orbits, around which there are secondary resonances, and so on [3]). Similarly, near a

homoclinic loop to a saddle center of a non-integrable system many stable periodic orbits co-exist [29, 28]. When the corresponding periodic paths in the configuration space do not intersect, one can obtain KAM-stable motions of repelling particles along several such paths: the particles on the same path must have the same frequency to avoid collisions, but the particles on different paths may have different frequencies. When the paths do intersect, one needs the frequencies of the motion along the different paths to be in resonance. This condition is not actually restrictive: given any finite set of elliptic orbits with arbitrary periods, one can tune the partial energies such that for every two paths the ratio of the periods would become rational. For particles in the billiard this is done just by normalizing the motion speed for each path by the path length.

Choreographies in the box. The billiard in the box is integrable, and therefore we build choreographic multi-particle regimes based on families of parabolic periodic solutions (instead of elliptic orbits). We have considered only one of such families in this paper, of parallel vertical motion, but there are many types of them, for example, diamond shaped orbits. In our construction of Section 5, to avoid collisions, a single particle sits on each periodic path. However, for other types of parabolic families, such as diamond shape orbits in a $d \geq 3$ box, many particles may occupy the same path and parallel paths without any collisions. In particular, it may be possible to create KAM-stable choreographic motions of $N \propto (l/\rho)^3$ particles in a three-dimensional box of characteristic size l . Similar solutions can be built for particles in ellipsoidal billiard, and for other systems of weakly-interacting particles whose individual dynamics is integrable.

Dynamics of the averaged system. We have shown that the stable choreographic motions are controlled by effective potentials defined on the torus corresponding to the set of phases of the individual particles. Our main result only refers to the fact that near the minimum of such potential the dynamics are, generically, KAM stable. However, one may also ask a question of the global dynamics: can the averaged system be completely integrable, or can one find additional stable motions far from the minimum of the potential? Various types of KAM-stable solutions of the averaged system should generate new types of non-trivial coherent states which may depend differently on physical parameters and be relevant for a larger variety of physical settings than the dynamically simplest types of choreographies we found here. Also, as explained in Appendix A, in the particular case of the equidistant particles' phases, our effective potentials share the same symmetries as the potentials of the classical Fermi-Pasta-Ulam chains, yet they form a larger class, and it may be interesting to study this broader class of systems.

Solid coherent states in a high temperature gas. In this paper, we operate in the limit where the motion along the periodic orbit is faster than the oscillations of the phase

differences between the particles. For highly energetic particles in a container, if the non-averaged interaction potential has a minimum (like the Lennard-Jones potential), one can think of an opposite limit where the frequencies ω of the small oscillations of the particles near this minimum are much faster than the frequency of the periodic billiard-like motion of the center of mass: $\omega \gg \frac{\sqrt{2h}}{L^*}$ (the kinetic energy of the center-of-mass motion can still be much larger than the energy of the fast but small oscillations of the particles: $I \ll h/\omega$). It may be interesting to study whether this can also lead to stable choreographic motions of molecules.

Large N limit. As one can see, different types of coherent states can correspond to different scaling of the number of particles as a function of the size of the container. In general, when considering the limit $N \rightarrow \infty$, one should also decide how the parameters of the system (the system size l , the energy per particle h , effective particle diameter ρ , etc.) scale with N : different types of scalings correspond to different physical situations. Determining which type of scalings correspond to various stable coherent states is the key for resolving the question of the realizability of such states in physically relevant settings.

Acknowledgements:

VRK work was funded by Israel Science Foundation (grant 787/22). DT work was funded by the Leverhulme Trust (grant RPG-2021-072). We thank V. Gelfreich and L. Chierchia for useful discussions.

Appendix

A Choreographic solutions with equidistant phases.

The evolution of the phases of the choreographic solutions is described, in the first-order approximation, by the averaged Hamiltonian

$$H = \frac{1}{2}(I_0)^2 + U(\theta), \quad (\theta, I_0) \in \mathbb{T}^N \times \mathbb{R}^N, \quad (183)$$

see Lemma 3.1 (here we scale I_0 such that $a = 1$ in the matrix A of (46), see (7)). The averaged potential U is given by

$$U(\theta^{(1)}, \dots, \theta^{(N)}) = \sum_{\substack{n,m=1,\dots,N \\ n \neq m}} W_{avg}(\theta^{(n)} - \theta^{(m)}),$$

where W_{avg} is an even, 2π -periodic function defined by (11).

System (183) is a generalization of the classical Fermi-Pasta-Ulam chain: the difference is that in the FPU we have only $m = n + 1 \pmod N$ in the sum describing the potential U , i.e., the interaction between the phases θ is short-range, while in our case all phases typically interact with each other. Like in the FPU, any uniformly distributed particle configuration $\theta = \theta_{eq}$, where

$$\theta_{eq}^{(2)} - \theta_{eq}^{(1)} = \dots = \theta_{eq}^{(N)} - \theta_{eq}^{(N-1)} = \theta_{eq}^{(1)} + 2\pi - \theta_{eq}^{(N)} = \frac{2\pi}{N}$$

is an equilibrium of system (183). Indeed, it is always an extremum of U since $\frac{\partial U}{\partial \theta^{(n)}}|_{\theta_{eq}} = 0$ for each n :

$$\frac{\partial U}{\partial \theta^{(n)}}|_{\theta_{eq}} = \sum_{m \neq n} W'_{avg}(\theta_{eq}^{(n)} - \theta_{eq}^{(m)}) = \sum_{k=1}^{N-1} W'_{avg}\left(\frac{2\pi}{N}k\right) = 0,$$

where the last equality follows because W'_{avg} is an odd and 2π -periodic function.

Such configurations form a line of extrema of U (parameterized by the choice of $\theta_{eq}^{(1)}$). It is a line of *minima* of U when the Hessian matrix at $\theta = \theta_{eq}$ is positive semi-definite. We have

$$\frac{\partial^2 U}{\partial \theta^{(n)} \partial \theta^{(m)}}(\theta_{eq}) = \begin{cases} -W''_{avg}\left(\frac{2\pi(n-m)}{N}\right) = u_{n-m} & \text{for } n \neq m \\ \sum_{j=1}^{N-1} W''_{avg}\left(\frac{2\pi j}{N}\right) = u_0 = -\sum_{k=1}^{N-1} u_k & \text{for } n = m. \end{cases} \quad (184)$$

Note that the numbers u_k satisfy $u_k = u_{-k} = u_{N-k}$ because W''_{avg} is even and 2π -periodic.

The Hessian matrix (184) is a circulant matrix, so its eigenvectors are the Fourier modes

$$v_j = \left(1, \exp\left(i\frac{2\pi j}{N}\right), \exp\left(i\frac{4\pi j}{N}\right), \dots, \exp\left(i\frac{2(N-1)j\pi}{N}\right)\right), \quad j = 0, \dots, N-1.$$

The corresponding eigenvalues are:

$$\lambda_j = u_0 + u_1 \exp\left(i\frac{2\pi j}{N}\right) + u_2 \exp\left(i\frac{4\pi j}{N}\right) + \dots + u_{N-2} \exp\left(i\frac{2\pi j(N-2)}{N}\right) + u_{N-1} \exp\left(i\frac{2\pi j(N-1)}{N}\right),$$

i.e.,

$$\lambda_j = 2 \sum_{1 \leq k < N/2} u_k \left(\cos\left(k\frac{2\pi j}{N}\right) - 1\right) + ((-1)^j - 1)u_{N/2} \quad (185)$$

(we take $u_{N/2} = 0$ in this formula when N is odd).

As we see, $\lambda_0 = 0$, which is due to the translational symmetry of the Hamiltonian (183). When all other λ_j are strictly positive, the line of uniformly distributed particles' configurations consists of minima of the potential U . This happens, for example, when

$u_k < 0$ for all $k \neq 0$, i.e., when W_{avg} is a convex function, which is consistent with the repelling nature of the interaction.

The frequencies of small oscillations around the line of minima are equal to $\sqrt{\lambda_j}$. It follows from (185) that

$$\lambda_j = \lambda_{N-j}$$

for all $1 \leq j \leq N/2$. Hence, the standard non-resonance assumption on the interaction potential breaks at $\theta = \theta_{eq}$. In particular, adding a small perturbation to (183) (without breaking the translational symmetry) would result, in general, in the destruction of the ellipticity.

Yet, we show next that the KAM Assumption IP1 still holds generically. Notice that the resonance relations are due to the discrete symmetries of the potential near $\theta = \theta_{eq}$. Namely, the system in a small neighborhood of the minima line is symmetric with respect to the transformations $T : \theta^{(n)} \rightarrow \theta^{(n+1)}$ (where n is taken mod N) and $S : \theta^{(n)} \rightarrow \theta^{(N+1-n)}$. The maps S and T generate the so-called N -th dihedral group. This is the same group of symmetries as in the FPU. The normal form theory for the FPU was built by Rink [39]. In fact, he derived the normal form near an equilibrium of a general S, T -symmetric Hamiltonian provided the Hamiltonian has no additional resonances. His work applies to our Hamiltonian (even though the interactions here are for all particle pairs whereas in [39] the FPU chain with only nearest neighbors interactions was considered). Indeed, notice that for any prescribed sequence $u_k, k = 1, \dots, N-1$ there is a smooth even and 2π -periodic potential W_{avg} satisfying $W''_{avg}(\frac{2\pi}{N}k) = u_k$, so, generically, for our system, no additional independent resonance relations appear.

The symmetric Rink normal form is given by Theorem 8.2 in [39] and for *odd* N is written as follows:

$$H_{Rink} = \sum_{1 \leq j < \frac{N}{2}} \sqrt{\lambda_j} a_j + \frac{1}{2} \sum_{1 \leq j, k < \frac{N}{2}} (C_{jk}^a a_j a_k + C_{jk}^b b_j b_k) \quad (186)$$

whereas for *even* N

$$H_{Rink} = \sum_{1 \leq j \leq \frac{N}{2}} \sqrt{\lambda_j} a_j + \frac{1}{2} \sum_{1 \leq j, k \leq \frac{N}{2}} C_{jk}^a a_j a_k + \frac{1}{2} \sum_{1 \leq j, k < \frac{N}{2}} C_{jk}^b b_j b_k + \sum_{1 \leq j \leq \frac{N}{4}} C_j^d (d_j d_{\frac{N}{2}-j} - c_j c_{\frac{N}{2}-j}). \quad (187)$$

The terms a, b, c, d are quadratic functions, so the normal form is of order 4 (the C 's are constant coefficients such that $C_{jk}^* = C_{kj}^*, * \in \{a, b, c, d\}$). The transition to the normal form is done as follows. First, one makes a linear symplectic coordinate transformation which

diagonalizes the quadratic part of the Hamiltonian near $(\theta = \theta_{eq}, I_0 = 0)$. One defines

for $1 \leq j < N/2$:

$$\begin{aligned} z_j &= \frac{1}{\sqrt{2N}} \sum_{n=1}^N e^{-2\pi i \frac{j}{N} n} (I_0^{(n)} + i\sqrt{\lambda_j} (\theta^{(n)} - \theta_{eq}^{(n)})), \\ \zeta_j &= \frac{1}{\sqrt{2N\lambda_j}} \sum_{n=1}^N e^{2\pi i \frac{j}{N} n} (I_0^{(n)} - i\sqrt{\lambda_j} (\theta^{(n)} - \theta_{eq}^{(n)})), \\ z_{N-j} &= -\frac{1}{\sqrt{2N}} \sum_{n=1}^N e^{-2\pi i \frac{j}{N} n} (I_0^{(n)} - i\sqrt{\lambda_j} (\theta^{(n)} - \theta_{eq}^{(n)})), \\ \zeta_{N-j} &= \frac{1}{\sqrt{2N\lambda_j}} \sum_{n=1}^N e^{2\pi i \frac{j}{N} n} (I_0^{(n)} + i\sqrt{\lambda_j} (\theta^{(n)} - \theta_{eq}^{(n)})), \end{aligned}$$

and, if N is even:

$$\begin{aligned} z_{\frac{N}{2}} &= \frac{1}{\sqrt{2N\lambda_{\frac{N}{2}}}} \sum_{n=1}^N (-1)^n (I_0^{(n)} + i\sqrt{\lambda_j} (\theta^{(n)} - \theta_{eq}^{(n)})), \\ \zeta_{\frac{N}{2}} &= -\frac{1}{\sqrt{2N}} \sum_{n=1}^N (-1)^n (I_0^{(n)} - i\sqrt{\lambda_j} (\theta^{(n)} - \theta_{eq}^{(n)})), \end{aligned}$$

cf. [39], formulas (7.2),(7.3). By the translational invariance of the average potential U , the Hamiltonian (183) in these coordinates is the sum of the term $\frac{1}{2N} \sum_{n=1}^N (I_0^{(n)})^2$ and a function which depends only on (z_j, ζ_j) with $j = 1, \dots, N-1$. This function is the Hamiltonian of the system reduced by the translation symmetry group $(\theta^{(n)} \mapsto \theta^{(n)} + c, c \in \mathbb{R}^1, n = 1, \dots, N)$. Since $(z, \zeta) = 0$ is an equilibrium of the reduced system, the Taylor expansion of the reduced Hamiltonian at zero starts with quadratic terms – these terms coincide with the quadratic part of H_{Rink} in (186) and (187), i.e., with $\sum_{1 \leq j \leq \frac{N}{2}} \sqrt{\lambda_j} a_j$. Next, one does a symplectic transformation (identity plus terms of the second order and higher) of the variables (z_j, ζ_j) , $j = 1, \dots, N-1$, which brings the reduced Hamiltonian to the form which coincides with the normal forms (186) or (187) up to terms of order 5. It is a standard fact that this system has KAM-tori if the normal form has KAM-tori in an arbitrarily small neighborhood of zero, so we further focus on finding KAM-tori in the normal forms.

In the new variables (z, ζ) the functions a, b, c, d in H_{Rink} are given by

$$\begin{aligned} a_j &= i(z_j \zeta_j - z_{N-j} \zeta_{N-j}), & b_j &= i(z_j \zeta_j + z_{N-j} \zeta_{N-j}), \\ c_j &= \frac{1}{\sqrt{\lambda_j}} (z_j z_{N-j} + \lambda_j \zeta_j \zeta_{N-j}), & d_j &= \frac{-i}{\sqrt{\lambda_j}} (z_j z_{N-j} - \lambda_j \zeta_j \zeta_{N-j}), \end{aligned}$$

for $1 \leq j < \frac{N}{2}$; we also have, for even N ,

$$a_{\frac{N}{2}} = iz_{\frac{N}{2}} \zeta_{\frac{N}{2}}.$$

Note that a, b, c, d are real and satisfy

$$a_j^2 = b_j^2 + c_j^2 + d_j^2, \quad (188)$$

see [39], formulas (8.6),(8.7). The coefficients C_{jk}^* in H_{Rink} are polynomials of the coefficients of the original Hamiltonian depending rationally on $\sqrt{\lambda_j}$. One can check that a generic potential U corresponds to a generic choice of C 's.

For odd N , the normal form (186) is completely integrable and generically has KAM tori [39]. This gives us that the the KAM assumption IP1 is fulfilled generically for system (183) for odd N . Let us consider the case of even N . By [39], Corollary 9.3, the $(N-1)$ degrees of freedom system (187) has a number of quadratic integrals in involution: $a_j, j = 1, \dots, \frac{N}{2}$ and $b_j - b_{\frac{N}{2}-j}, 1 \leq j < \frac{N}{4}$. This set is incomplete; while the normal form H_{Rink} for the nearest-neighbor FPU chain with an even N has additional integrals and is completely integrable [19, 20], it is not known whether this normal form is completely integrable for a generic choice of coefficients C and λ in (187), or for a general choice of the potential U in (183).

However, the restriction of (187) to the invariant subspace $\{a_s = 0, 1 \leq s \leq \frac{N}{4}\}$ is completely integrable: since b_s, c_s and d_s all vanish for $1 \leq s \leq \frac{N}{4}$ by (188), the restricted Hamiltonian is given by

$$H = \sum_{\frac{N}{4} < j \leq \frac{N}{2}} \sqrt{\lambda_j} a_j + \frac{1}{2} \sum_{\frac{N}{4} < j, k \leq \frac{N}{2}} C_{jk}^a a_j a_k + \frac{1}{2} \sum_{\frac{N}{4} < j, k < \frac{N}{2}} C_{jk}^b b_j b_k,$$

and the quadratic functions a_j ($\frac{N}{4} < j \leq \frac{N}{2}$) and b_j ($\frac{N}{4} < j < \frac{N}{2}$) give a complete set of its integrals. The restricted Hamiltonian has the same structure as the completely integrable normal form (186). So, in the same way as it was done in [39] for system (186), one establishes that for a generic choice of non-zero values of the integrals a_j ($\frac{N}{4} < j \leq \frac{N}{2}$) and b_j ($\frac{N}{4} < j < \frac{N}{2}$) the corresponding joint level set of these integrals is a KAM-torus \mathcal{T} of the restricted system. Moreover, these integrals a_j and b_j are the action variables (see formula (9.3) in [39]). The dynamics in a small neighborhood of such torus are described, in the main approximation, by the Hamiltonian (187) *averaged* over the angle variables conjugate to the actions. The only terms in (187) that depend on these angle variables are $d_{\frac{N}{2}-j}$ and $c_{\frac{N}{2}-j}$ with $j < \frac{N}{4}$; let us show that their averaged values are zero.

Indeed, by the ergodicity of the flow on the invariant torus \mathcal{T} , we can replace the averaging

over the angle variables by the time averaging. Due to the commutation relations (see formula (9.1) in [39])

$$\{b_k, c_m\} = 2d_m \delta_{km}, \quad \{b_k, d_m\} = -2d_m, \quad \{a_k, c_m\} = \{a_k, d_m\} = 0,$$

where $\{\cdot, \cdot\}$ is the Poisson bracket and δ_{km} is the Kronecker delta, we have that for $\frac{N}{4} < m < \frac{N}{2}$

$$\frac{d}{dt}c_m = \{c_m, H_{Rink}\} = -\Omega_m d_m, \quad \frac{d}{dt}d_m = \{d_m, H_{Rink}\} = \Omega_m c_m,$$

where

$$\Omega_m = 2 \sum_{\frac{N}{4} < k < \frac{N}{2}} C_{km}^b b_k.$$

For a generic choice of the actions b_k ($\frac{N}{4} < k < \frac{N}{2}$), the frequencies Ω_m are all non-zero, hence c_m and d_m perform harmonic oscillations and their time-average is zero. Thus, the averaged system (187) is

$$H = \sum_{1 \leq j \leq \frac{N}{2}} \sqrt{\lambda_j} a_j + \frac{1}{2} \sum_{1 \leq j, k \leq \frac{N}{2}} C_{jk}^a a_j a_k + \frac{1}{2} \sum_{1 \leq j, k < \frac{N}{2}} C_{jk}^b b_j b_k.$$

This system has the same structure as (186) and is completely integrable (the integrals are a_j and b_j). It follows that the invariant torus \mathcal{T} is normally elliptic. As in [39], one checks that a generic Liouville torus of this systems satisfies the twist condition, hence \mathcal{T} is surrounded by KAM-tori and these tori persist when we proceed from the averaged system to the original system H_{Rink} in a neighborhood of \mathcal{T} .

References

- [1] R. Abraham and J. E. Marsden. *Foundations of mechanics*. Number 364. American Mathematical Soc., 2008.
- [2] C. Antonopoulos, T. Bountis, and C. Skokos. Chaotic dynamics of n-degree of freedom hamiltonian systems. *International Journal of Bifurcation and Chaos*, 16(06):1777–1793, 2006.
- [3] V. Arnold, V. Kozlov, and A. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*, volume 3. Springer Science & Business Media, 2007.
- [4] P. Bálint, T. Gilbert, D. Szász, and I. P. Tóth. What mathematical billiards teach us about statistical physics? *Pure and Applied Functional Analysis*, 6(1):1–35, 2021.

- [5] L. Barreira and Y. Pesin. Smooth ergodic theory and nonuniformly hyperbolic dynamics. In *Handbook of dynamical systems. Vol. 1B*, pages 57–263. Elsevier B. V., Amsterdam, 2006. With an appendix by Omri Sarig.
- [6] L. Bunimovich. Kinematics, equilibrium, and shape in Hamiltonian systems: the “LAB” effect. *Chaos. An Interdisciplinary Journal of Nonlinear Science*, 13(3):903–912, 2003.
- [7] L. Bunimovich, C. Liverani, A. Pellegrinotti, and Y. Suhov. Ergodic systems of n balls in a billiard table. *Communications in Mathematical Physics*, 146(2):357–396, 1992.
- [8] L. A. Bunimovich, Y. G. Sinai, and N. I. Chernov. Statistical properties of two-dimensional hyperbolic billiards. *Russian Mathematical Surveys*, 46(4):47–106, aug 1991.
- [9] A. Chenciner, J. Gerver, R. Montgomery, and C. Simó. Simple choreographic motions of N bodies: a preliminary study. In *Geometry, mechanics, and dynamics*, pages 287–308. Springer, New York, 2002.
- [10] A. Chenciner and R. Montgomery. A remarkable periodic solution of the three-body problem in the case of equal masses. *Annals of Mathematics. Second Series*, 152(3):881–901, 2000.
- [11] N. Chernov and R. Markarian. *Chaotic Billiards*. Number 127. American Mathematical Soc., 2006.
- [12] L. Chierchia and G. Pinzari. Properly-degenerate kam theory (following vi arnold). *Discrete & Continuous Dynamical Systems-S*, 3(4):545, 2010.
- [13] L. Chierchia and G. Pinzari. The planetary n -body problem: symplectic foliation, reductions and invariant tori. *Inventiones mathematicae*, 186(1):1–77, 2011.
- [14] M. M. Dvorin and V. F. Lazutkin. Existence of an infinite number of elliptic and hyperbolic periodic trajectories for convex billiards. *Akademiya Nauk SSSR. Funkcionalnyi Analiz i ego Prilozenija*, 7(2):20–27, 1973.
- [15] J. Fejoz, A. Knauf, and R. Montgomery. Classical n -body scattering with long-range potentials. *Nonlinearity*, 34(11):8017–8054, 2021.
- [16] H. Fukuda, T. Fujiwara, and H. Ozaki. Figure-eight choreographies of the equal mass three-body problem with lennard-jones-type potentials. *Journal of Physics A: Mathematical and Theoretical*, 50(10):105202, feb 2017.

- [17] G. Fusco, G. F. Gronchi, and P. Negrini. Platonic polyhedra, topological constraints and periodic solutions of the classical n -body problem. *Inventiones mathematicae*, 185(2):283–332, 2011.
- [18] M. Guardia, P. Martin, J. Paradela, and T. M. Seara. Hyperbolic dynamics and oscillatory motions in the 3 body problem, 2022.
- [19] A. Henrici and T. Kappeler. Results on normal forms for FPU chains. *Communications in Mathematical Physics*, 278(1):145–177, 2008.
- [20] A. Henrici and T. Kappeler. Resonant normal form for even periodic FPU chains. *Journal of the European Mathematical Society (JEMS)*, 11(5):1025–1056, 2009.
- [21] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.
- [22] A. Knauf. The n -centre problem of celestial mechanics for large energies. *Journal of the European Mathematical Society (JEMS)*, 4(1):1–114, 2002.
- [23] V. Kozlov and D. Treschev. *A genetic introduction to the dynamics of systems with impacts*. AMS, Providence, 1991.
- [24] V. V. Kozlov. Billiards, invariant measures, and equilibrium thermodynamics. *Regular & Chaotic Dynamics. International Scientific Journal*, 5(2):129–138, 2000.
- [25] V. V. Kozlov. Polynomial conservation laws for the Lorentz and the Boltzmann-Gibbs gases. *Uspekhi Matematicheskikh Nauk*, 71(2(428)):81–120, 2016.
- [26] N. S. Krylov. *Works on the foundations of statistical physics*. Princeton Series in Physics. Princeton University Press, Princeton, N.J., 1979. Translated from the Russian by A. B. Migdal, Ya. G. Sinai [Ja. G. Sinaĭ] and Yu. L. Zeeman [Ju. L. Zeeman], With a preface by A. S. Wightman, With a biography of Krylov by V. A. Fock [V. A. Fok], With an introductory article “The views of N. S. Krylov on the foundations of statistical physics” by Migdal and Fok, With a supplementary article “Development of Krylov’s ideas” by Sinaĭ.
- [27] S. Lansel, M. A. Porter, and L. A. Bunimovich. One-particle and few-particle billiards. *Chaos. An Interdisciplinary Journal of Nonlinear Science*, 16(1):013129, 12, 2006.

- [28] L. Lerman. Hamiltonian systems with loops of a separatrix of a saddle-center. *Selecta Math. Sov.*, 10(3):297–306, 1991.
- [29] L. Lerman and K. Trifonov. Saddle-center and periodic orbit: Dynamics near symmetric heteroclinic connection. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(2):023113, 2021.
- [30] K. Meyer, G. Hall, and D. Offin. *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, volume 90. Springer Science & Business Media, 2008.
- [31] R. Montgomery. N-body choreographies. *Scholarpedia*, 5(11):10666, 2010. revision #137285.
- [32] R. Montgomery. The three-body problem and the shape sphere. *American Mathematical Monthly*, 122(4):299–321, 2015.
- [33] C. Moore. Braids in classical dynamics. *Physical Review Letters*, 70(24):3675–3679, 1993.
- [34] A. Neishtadt, K. Sheng, and V. Sidorenko. Stability analysis of apsidal alignment in double-averaged restricted elliptic three-body problem. *Celestial Mechanics and Dynamical Astronomy*, 133(10):1–23, 2021.
- [35] T. Ouyang and Z. Xie. Star pentagon and many stable choreographic solutions of the newtonian 4-body problem. *Physica D: Nonlinear Phenomena*, 307:61–76, 2015.
- [36] A. Rapoport and V. Rom-Kedar. Nonergodicity of the motion in three-dimensional steep repelling dispersing potentials. *Chaos*, 16(4):043108, 6, 2006.
- [37] A. Rapoport, V. Rom-Kedar, and D. Turaev. Approximating multi-dimensional Hamiltonian flows by billiards. *Comm. Math. Phys.*, 272(3):567–600, 2007.
- [38] A. Rapoport, V. Rom-Kedar, and D. Turaev. Stability in high dimensional steep repelling potentials. *Comm. Math. Phys.*, 279(2):497–534, 2008.
- [39] B. Rink. Symmetry and resonance in periodic fpu chains. *Communications in Mathematical Physics*, 218(3):665–685, 2001.
- [40] V. Rom-Kedar and D. Turaev. Billiards: A singular perturbation limit of smooth Hamiltonian flows. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(2):026102, 2012.

- [41] N. Simányi. The complete hyperbolicity of cylindric billiards. *Ergodic Theory and Dynamical Systems*, 22(1):281–302, 2002.
- [42] N. Simányi. Proof of the Boltzmann-Sinai ergodic hypothesis for typical hard disk systems. *Inventiones Mathematicae*, 154(1):123–178, 2003.
- [43] N. Simányi. Proof of the ergodic hypothesis for typical hard ball systems. *Annales Henri Poincaré. A Journal of Theoretical and Mathematical Physics*, 5(2):203–233, 2004.
- [44] N. Simányi and D. Szász. Hard ball systems are completely hyperbolic. *Annals of Mathematics. Second Series*, 149(1):35–96, 1999.
- [45] C. Simó. Dynamical properties of hamiltonian systems with applications to celestial mechanics. In *Central Configurations, Periodic Orbits, and Hamiltonian Systems*, pages 169–232. Springer, 2015.
- [46] Y. G. Sinai. On the foundations of the ergodic hypothesis for a dynamical system of statistical mechanics. In *Doklady Akademii Nauk*, volume 153, pages 1261–1264. Russian Academy of Sciences, 1963.
- [47] Y. G. Sinai. Dynamical systems with elastic reflections. *Russian Mathematical Surveys*, 25(2):137, 1970.
- [48] Y. G. Sinai. Hyperbolicity and chaos. In *Boltzmann’s legacy 150 years after his birth (Rome, 1994)*, volume 131 of *Atti Convegni Lincei*, pages 107–110. Accad. Naz. Lincei, Rome, 1997.
- [49] J. G. Sinaĭ. Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Akademiya Nauk SSSR i Moskovskoe Matematicheskoe Obshchestvo. Uspekhi Matematicheskikh Nauk*, 25(2 (152)):141–192, 1970.
- [50] D. Turaev. Richness of chaos in the absolute newhouse domain. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures*, pages 1804–1815. World Scientific, 2010.
- [51] D. Turaev and V. Rom-Kedar. Elliptic islands appearing in near-ergodic flows. *Nonlinearity*, 11(3):575, 1998.
- [52] D. Turaev and V. Rom-Kedar. Soft billiards with corners. *J. Statist. Phys.*, 112(3-4):765–813, 2003.

- [53] M. P. Wojtkowski. Design of hyperbolic billiards. *Communications in Mathematical Physics*, 273(2):283–304, 2007.
- [54] G. Yu. Simple choreographies of the planar newtonian n-body problem. *Archive for Rational Mechanics and Analysis*, 225(2):901–935, 2017.