CLASSICAL BILLIARDS AND QUANTUM LARGE DEVIATIONS

BY MARCO LENCI

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Joel L. Lebowitz

and approved by

New Brunswick, New Jersey

October, 1999

ABSTRACT OF THE DISSERTATION

Classical Billiards and Quantum Large Deviations

by Marco Lenci

Dissertation Director: Joel L. Lebowitz

This dissertation treats two topics in the general realm of probability in mathematical physics:

Part I, **Infinite step billiards**, deals with the classical motion in a certain family of non-compact billiards, defined via the following: To a given sequence of non-negative numbers $\{p_n\}_{n\in\mathbb{N}}$, such that $p_n \searrow 0$, there corresponds a billiard *table* $P := \bigcup_{n\in\mathbb{N}}[n, n+1] \times [0, p_n]$.

For these dynamical systems, we derive two categories of results. In the context of topological dynamics, we study the so-called *escape orbits*. In particular, we find that, for a large set of infinite step billiards, there is almost surely a unique escape orbit which is somehow an attractor for every other trajectory. From the viewpoint of strict ergodic theory, we present results about the existence of ergodic billiards. The main theorem states that *generically* these systems are ergodic for almost all initial velocities.

In Part II, **Large deviations for ideal quantum systems**, we consider a general d-dimensional quantum system of non-interacting particles in a very large (formally infinite) container.

We prove that, in equilibrium, the fluctuations in the density of particles in a subdomain Λ of the container are described by a large deviation function related to the

pressure of the system. That is, untypical densities occur with a probability exponentially small in the volume of Λ, with the coefficient in the exponent given by the appropriate thermodynamic potential. Furthermore, small fluctuations satisfy the central limit theorem.

Acknowledgements

I am very grateful to my advisor, Joel Lebowitz, for his support during the preparation of this thesis and, in general, in the course of my graduate studies at Rutgers. He was constantly interested in my progress and always found some time for me off his busy schedule.

I also wish to thank Sheldon Goldstein, Mikhail Lyubich and Eugene Speer for serving in my defense committee and for useful comments. Special thanks to Michael Kiessling with whom I had a few beneficial discussions.

Mirko Degli Esposti, Gianluigi Del Magno, Herbert Spohn and Joel, of course, collaborated on several results contained in the dissertation, which would have been very meager without their contribution!

I cannot help but acknowledge also the two mentors that I had before coming to Rutgers. I am tremendously indebted to Sandro Graffi and Yakov Sinai, who have given me enthusiasm for this job, peerless guidance, and constant help along the way.

If I made it through Graduate School, this is due to the heartening presence of a bunch of great friends, rather than my mathematical skills. I would like to mention (in order of appearance): Terri, Emilio, Jason, Chuck, the volleyball clique, Luca, Ryan, Bala, and Yuka. Stefano and Carlo provided precious technical support during the preparation of my defense talk.

Last but not least, deep thanks to my family.

Mom, I promise, no more dissertations!

Dedication

To all those who want to get the hell out of Graduate School.

Table of Contents

List of Figures

Part I

Infinite step billiards

Chapter 1

Introduction

The theory of dynamical systems is the tool that mathematicians have devised to study the time evolution of systems of physical interest. The first basic fact one learns as a student is that the knowledge of the equations that govern this evolution does not ensure that one can in practice predict the state of the system for future times.

This "general principle" is as old as mechanics itself, but only in relatively recent times has it become common knowledge that this *unpredictability* is not (only) a consequence of the complicated nature of the dynamical equations, or of the large number of degrees of freedom. Systems with very few degrees of freedom and very simple evolution laws can exhibit *chaotic* behavior. Although the solutions to the equations are unique—and thus completely determined by the choice of the initial conditions—the dependence on the initial conditions does not allow for approximations that hold for a long time. This phenomenon goes by the name of *deterministic chaos*.

Only for a very few systems are mathematicians able to integrate the equations of motion, whence the term *integrable* systems. (It is hard to give a precise definition of "very few", especially after the discovery of the KAM theorem, but everybody who has come across the field takes that as "none but the trivial examples in the textbooks".) For all other systems, the theory actually tries to deal with their *intrinsic* unpredictability and one of the main tools it employs is probability. This branch of the theory of dynamical systems is called *ergodic theory*.

One of the tasks ergodic theory takes on is that of classifying systems according to their chaotic properties. For instance, in vague terms, a dynamical system is *ergodic* when there are no non-trivial constants of the motion; *chaotic* when there is sensitive dependence on the initial conditions (this is not a well-established notion, but rather chaoticity encompasses many mathematical definitions such as *mixing*, *Anosov systems*, *K-systems* and others). We will use some of these definitions in this part of the work, assuming the reader is well familiar with them. Standard references include [AA, CFS, Ma, W].

1.1 Billiards

We have said that the vast majority of the systems in which physicists are interested cannot be integrated. As a matter of fact, they are hard to study from the point of view of ergodic theory, too. Therefore mathematicians are forced to consider models, abstract simplified systems that are easier to approach but, one hopes, still retain the relevant features of the physical systems they represent.

Billiards may seem very far from anything a mathematical physicist would come across during his or her working hours. However, billiard models have been extremely popular in the last decades.

From a mathematical point of view, a billiard is a dynamical system defined by the uniform motion of a point inside a domain (called the *table*) with *totally elastic* reflections at the boundary, i.e., reflections for which the tangential component of the velocity remains constant and the normal component changes sign.

The nice characteristics of these models is obviously that, for most of the time, they obey the simplest possible law of motion, that is, constant velocity. The complications arise only from the collisions with the boundary of the table. Hence, for these systems, geometric intuition can play a decisive role in making them treatable. On the other hand, billiards are known to exhibit the entire array of ergodic properties we have mentioned. For instance, billiards in rectangles and ellipses are integrable; billiards in generic polygons are ergodic, but not mixing; billiards in domains with dispersing boundaries can be K-systems (like the celebrated Sinai billiard [CFS]). A nice and complete reference is [T].

Other than the reason outlined above, the popularity of billiards is due to their wide use in the field of *quantum chaos* [Gu, H]. Certainly we do not want to delve into this subject here; let it suffice to say that research in quantum chaos tries to link the chaotic properties of classical Hamiltonian systems with the spectral properties of the Schrödinger operators in their quantum versions. For billiards, this reduces to studying the eigenvalue problem for the Laplacian on a certain domain, with Dirichlet boundary conditions. This is a problem that has a long tradition, so one need not start from scratch.

1.2 Billiards in polygons

Even in the case—apparently simple—of a bounded polygonal billiard (with a finite number of sides), it took several decades and some non-trivial techniques (see in particular the milestone work [KMS]) to produce what might be called a theory, a good treatment of which can be found in [G2, G3]. We give here a brief and incomplete review of some of the main results.

Usually one assumes that the magnitude of the particle's velocity equals one. Thus, if P is the table, the *phase space* is $P \times S^1$, endowed with the Lebesgue measure on each component. It is also commonly assumed that the trajectory which hits a vertex stops there. In fact, there is no general way to uniquely continue an orbit that encounters a vertex (a *singular orbit*, in jargon). However, the set of initial conditions which give rise to non-singular trajectories has always full measure. (It might be worthwhile to anticipate here that, for our models, there is actually a way to define even singular orbits for all times. We will adopt it in Section 2.1.)

Among the class of polygonal billiards, a table P is *rational* if the angles between the sides of P are all of the form $\pi n_i/m_i$, where the integers n_i and m_i are assumed to be the lowest terms of the fraction. In this case, a trajectory will have only a finite number of possible velocities. This implies a decomposition of the phase space in a family of flow-invariant surfaces S_θ, $0 \le \theta \le \pi/m$, $m := \text{l.c.m.}\{m_i\}$. If $\theta = 0, \pi/m$, then $S_\theta = P \times \{0, 2\pi/m, \ldots, 2\pi(m-1)/m\}$. In all other cases, $S_\theta = P \times \{\theta, \theta+\pi/m, \ldots, \theta+\pi(2m-1)/m\}$. These surfaces were first named (in [FK]) *Überlagerungsflächen*, but in this thesis we will spare the reader that term. There is a natural way to extend the Lebesgue measure (or any measure, for that matter) from P to S_θ. A planar representations of S_θ is obtained by the usual *unfolding* procedure for the orbits [FK, ZK].

Excluding the particular cases $\theta = 0, \pi/m$, it is known that the billiard flow restricted to any of the S_θ is essentially equivalent to a geodesic flow $\phi_{\theta,t}$ on a closed oriented surface S, endowed with a flat Riemannian metric with *conical* singularities. This surface is tiled by $2m$ copies of P, properly "glued" together. The topological type of the surface S, i.e., its genus g, is determined by the geometry of the rational polygon:

$$g(S) = 1 + \frac{m}{2} \sum_{i=1}^{n} \frac{n_i - 1}{m_i}. \tag{1.1}$$

With the use of this equivalence, a number of theorems regarding the existence and the number of ergodic invariant measures for the flow have been proven (cf. [ZK] and references therein).

More refined results concerning the billiard flows were then obtained by exploring the analogies of these flows with the *interval exchange transformations* (using the induced map on the boundary) on one hand [K, BKM, M, V1, V2], and with *holomorphic quadratic differentials* on compact Riemann surfaces, on the other [KMS, G3].

The deep connections between these three different subjects have been very useful in the understanding of polygonal billiard flows. We summarize some of the most important statements about rational polygons in the next proposition. (Briefly, let us recall that an *almost integrable billiard* is a billiard whose table is a finite connected union of pieces belonging to a tiling of the plane by reflection, e.g, a rectangular tiling, or a tiling by equilateral triangles, etc.)

Proposition 1.1 *The following statements hold true:*

(i) [KMS, Ar] *Given a (finite) rational polygon P and the family of its invariant surfaces $\{S_\theta\}$, the Lebesgue measure on S_θ is the unique ergodic measure for the billiard flow, for almost all θ.*

(ii) [ZK] *For all but countably many directions, a rational polygonal billiard is* minimal *(i.e., all infinite semi-orbits are dense) in S_θ.*

(iii) [G1, G2, B] *For almost integrable billiards, "minimal directions" and "ergodic directions" coincide.*

(iv) [GK] *Let* $\mathcal{R}_n$ *be the space of* n*-gons such that their sides are either horizontal or vertical, parametrized by the length of the sides. Then for any direction* θ, $0 < \theta < \pi/2$*, there is a dense* G_δ*-set in* $\mathcal{R}_n$*, such that for each polygon of this set the corresponding flow* $\phi_{\theta,t}$ *is weakly mixing.*

(v) [Ka] *For any rational polygon and any direction* θ*, the billiard flow* $\phi_{\theta,t}$ *is not mixing.*

Other important results, concerning generic polygonal tables, can be proven (we still refer to [G3] for a more exhaustive review):

Proposition 1.2 *The following statements hold true:*

(i) [G3, GKT] *For any given polygon, the metric entropy with respect to any flow-invariant measure is zero; furthermore, the topological entropy is also zero.*

(ii) [GKT] *Given an arbitrary polygon and an orbit, either the orbit is periodic or its closure contains at least one vertex.*

In Proposition 1.1,*(iv)* we have encountered the concept of a G_δ-subset of a topological space as a way to define a *typical* subset. This notion of typicality was first used by [ZK], and has become customary since [KMS]. Up to an affine transformation and a rescaling, every polygon with n sides is equivalent to a polygon with a vertex at the origin and an adjacent vertex at $(1, 0)$. The space of n-gons can thus be identified with some quotient of $\mathbb{R}^{2(n-2)}$. This is a complete metric space and the above definition applies.

It is very hard to establish ergodic properties for generic polygons, but one can use the results for rational polygons to approximate the former with the latter. Necessarily one ends up with density statements:

Proposition 1.3 *In the space of* n*-gons:*

(i) [ZK] *The typical polygon is* topologically transitive.

(ii) [KMS] *The typical polygon is ergodic for almost all directions θ.*

1.3 Non-compact billiards

Although the open questions about billiards in finite polygons are still many, it seems that the state-of-the-art could be considered fairly satisfactory, at least in terms of generic properties. However, until recently (before [DDL1] and especially [Tr]), even the simplest among the statements of Propositions 1.1-1.3 were unproven for (possibly non-compact) polygons with an infinite number of sides.

The interest in extending the previous results to infinite polygons goes beyond mere scientific curiosity. The billiard flow in a polygon with infinitely many sides is associated to an interval exchange transformation with a countable partition, which is a quite general automorphism of a measure-space. An intriguing example of a billiard in an infinite polygon, related to the models we will study henceforth, is provided by the "staircase" table briefly mentioned in [EFV].

Another motivation to pursue a generalization comes from interest in non-compact billiards. These are *open* systems (i.e., the particle might "escape at infinity") and one would like to see how this affects the dynamics. Even for a generic (non-necessarily polygonal) unbounded billiard, not much more is known than the existence (or non-existence) of certain types of trajectories [L, Ki, Le]. Furthermore, links between these questions and quantum chaos can be found in [Le]. Actually, since that article presents many analogies with this work (including the author), we will report its findings in Appendix A.

Some time ago Del Magno, Degli Esposti and Lenci [DDL1] introduced a family of billiards in certain non-compact infinite polygons. They were named *infinite step billiards*. An example is depicted in Fig. 1.1. This has incited Troubetzkoy to fill some voids in the area, extending some of the statements of Propositions 1.1-1.3 to the case of infinite polygons. His techniques have been recently adapted to the infinite step billiards in [DDL2], which also contains results that are more specifically related to these models. Further work is currently in progress [Tr2].

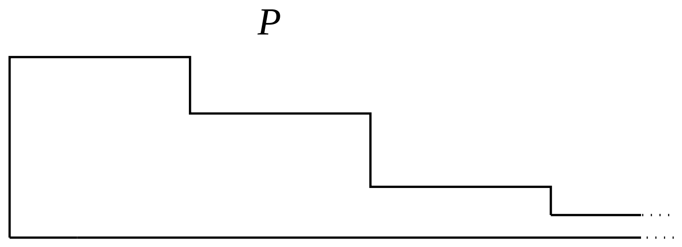

Figure 1.1: An infinite step billiard.

Infinite step billiards constitute the contents of Part I of this dissertation. In Chapter 2 we introduce the mathematical framework, including the fundamental concept of *escape orbit*. The analysis of the escape orbits turns out (rather unexpectedly) to have several implications for the dynamics of the billiards. We prove that, for almost all choices of the initial velocity, there is a unique escape orbit. This orbit is a topologically complex object (possibly a Cantor set) and acts somehow as an attractor for all other trajectories. The above is described in Chapter 3, for a large class of models, as proved in [DDL2]. However, the same results were already known in [DDL1] for a specific example. That model, the *exponential step billiard*, was studied in some detail, and is here presented in Chapter 4. Finally, Chapter 5 (that, once again, comes from [DDL2]) deals with the ergodic properties of our class of billiards, using the ideas of [Tr]. It is shown that, defining the space of infinite step billiards by means of a very natural metric, the typical table is ergodic for almost all initial velocities. Part I is completed by Section 5.3 which contains some results on the metric entropy.

Chapter 2

Infinite step billiards

An infinite step billiard is defined as follows: Consider a sequence $\{p_n\}_{n\in\mathbb{N}}$ of non-negative numbers such that $p_{n+1} \le p_n \; \forall n$ and $\lim_{n\to\infty} p_n = 0$. In words, $\{p_n\}$ is decreasing and vanishing. Also, let us fix $p_0 = 1$ (although this condition is inessential and will be relaxed in Chapter 5). The billiard table is the *step polygon* $P := \bigcup_{n\in\mathbb{N}}[n, n+1] \times [0, p_n]$ (Fig. 1.1), which is called *infinite* if $p_n > 0 \; \forall n$. Let (x, y) be the coordinates on P.

Following the considerations of Section 1.2, we see that a point particle can travel within P only in four directions $\theta \in S^1$ (two if the motion is vertical or horizontal—degenerate cases which we disregard). One of these directions lies in the first quadrant, $\theta \in]0, \pi/2[$. In the remainder, with the exception of Section 5.2, we will use the notation $\alpha := \tan\theta \in]0, +\infty[$. The invariant surface associated to the billiard flow is labeled by S^P_α (or just by S_α, when there is no means of confusion). It can be represented on the plane as in Fig. 2.1, with four copies of P, via the aforementioned unfolding procedure.

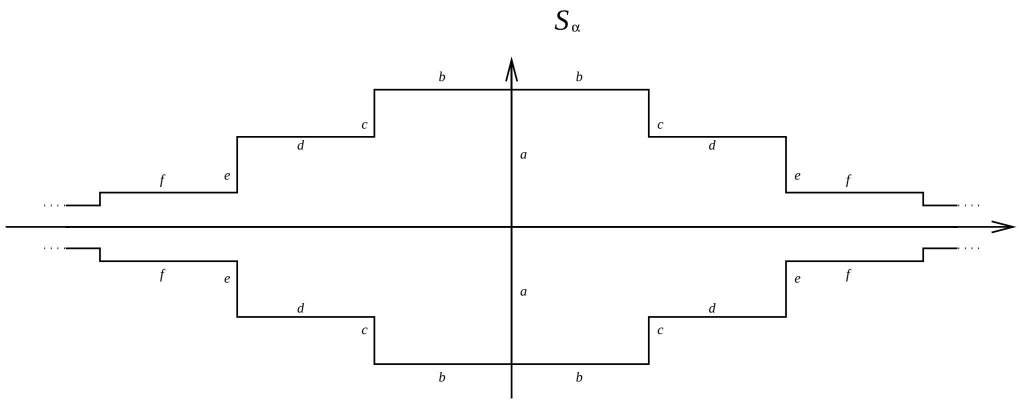

Figure 2.1: The invariant surface S^P_α for the infinite billiard.

We explain how to use Fig. 2.1: A point in S_α moves with constant speed along a line of slope α, heading upwards. Whenever it touches the boundary—say at some point on a horizontal side labeled by b—it reappears at the corresponding point on the lower side labeled by b. Then it continues its motion on another line of slope α. Analogously, when it hits a vertical side on the right, it reappears on the associated side on the left. A trajectory in S_α, then, is the union of segments of slope α, whose endpoints are on the boundaries of that domain. The intersection of S_α with the i-th quadrant is a copy of P (corresponding to the direction $\theta + i\pi/2$ for the velocity) and thus there is a natural projection $S_\alpha \longrightarrow P$. It is easy to see that the projection of a trajectory in S_α is a trajectory in P.

We denote by (X, Y) the intrinsic coordinates on S_α, inherited by the planar representation as in Fig. 2.1. There, the $3\pi/2$ corners represent the *non-removable singularities*, or *singular vertices*, V_k, of coordinates (k, p_k) on P, or $(\pm k, \pm p_k)$ on S. More precisely, recalling Section 1.2, V_k is a conical singularity. In fact, one sees that, in the metric induced by the Euclidean metric on the plane, it takes an angle 6π to go around it.

With the additional condition $\sum_n p_n < \infty$, S_α^P can be considered a non-compact, finite-area surface of infinite genus. We denote by $P^{(n)}$ the truncated billiard that one obtains by closing the table at $x = n$. The corresponding invariant surface is denoted by $S_\alpha^{P^{(n)}}$, or simply by $S_\alpha^{(n)}$ (Fig. 2.2), and (1.1) shows that it has genus n. Only to $S_\alpha^{(n)}$ can we apply the many statements of Section 1.2. The purpose of this work is precisely to extend some of those results to the infinite table P.

2.1 The return map

On S_α, the first vertical side of P becomes the closed curve $L := \{0\} \times [-1, 1[$ (keeping in mind that $(0, -1)$ and $(0, 1)$ are identified in Fig. 2.1) which separates S_α in two symmetric parts. We will occasionally identify L with the interval $[-1, 1[$.

The billiard flow along a direction α is denoted by $\phi_{\alpha,t}$ (or ϕ_t when there is no danger of confusion). It will be shown in the next section that, for $\alpha \neq +\infty$, the

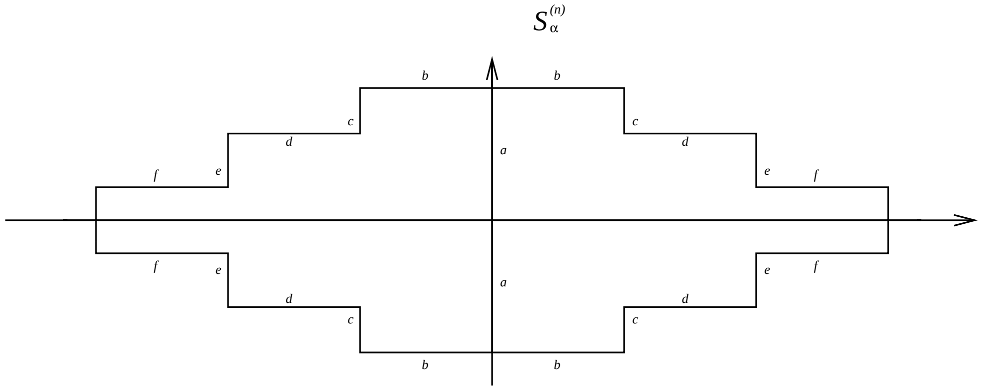

Figure 2.2: The invariant surface $S_\alpha^{(n)}$ for the truncated billiard.

forward semi-orbit of almost every point of L (w.r.t. the Lebesgue measure) intersects L again. This means that the flow induces a.e. on L a Poincaré map $\mathcal{P}_\alpha$, which we call the *(first) return map.* It is easy to see that $\mathcal{P}_\alpha$ is an infinite-partition *interval exchange transformation* (i.e.t.).

On L we establish the convention that the map is continuous from above: i.e., an orbit going to the singular vertex (n, p_n) of S_α will continue from the point $(-n, p_n)$, thus behaving like the orbits above it, i.e., bouncing backwards. In the same spirit, a trajectory hitting $(-n, p_n)$ will continue from $(-n, -p_n)$, while orbits encountering vertex $(n, -p_n)$ will just pass through. This corresponds to partitioning $L \simeq [-1, 1[$ into right-open subintervals.

The fact that the number of subintervals is infinite is exactly what makes the study of the ergodic properties of this system a non-trivial task.

It is now natural to relate $\mathcal{P}_\alpha$ to the family of return maps $\mathcal{P}_\alpha^{(n)}$ corresponding to the truncated billiards $P^{(n)}$. These are finite-partition i.e.t.'s defined on all of L (with abuse of notation, L also denotes the obvious closed curve on $S_\alpha^{(n)}$, Fig. 2.2).

2.2 Escape orbits

Let $E_\alpha^{(n)} \subset L$ be the set of points whose forward orbit starts along the direction α and reaches the n-th *aperture* $G_n := \{n\} \times [-p_n, p_n[$ without colliding with any vertical walls. $E_\alpha^{(n)}$ is the union of at most n right-open intervals, since the backward evolution of G_n can only *split* once for each of the $n-1$ singular vertices (Fig. 2.3). Hence $E_\alpha^{(n)} = \bigcup_{j=1}^{m_n} I_{\alpha,j}^{(n)}$. We denote $n.i.(E_\alpha^{(n)}) := m_n \le n$, where $n.i.$ stands for "number of intervals". Moreover, $|E_\alpha^{(n)}| = 2p_n$ and $E_\alpha^{(n+1)} \subset E_\alpha^{(n)}$. From this we infer that the family $\{I_{\alpha,j}^{(n)}\}_{n,j}$ can be rearranged into families of nested right-open intervals, whose lengths vanish as $n \to \infty$. Also, the sequence of i.e.t.'s $\mathcal{P}_\alpha^{(n)}$ converges a.e. in L to $\mathcal{P}_\alpha$, as $n \to \infty$.

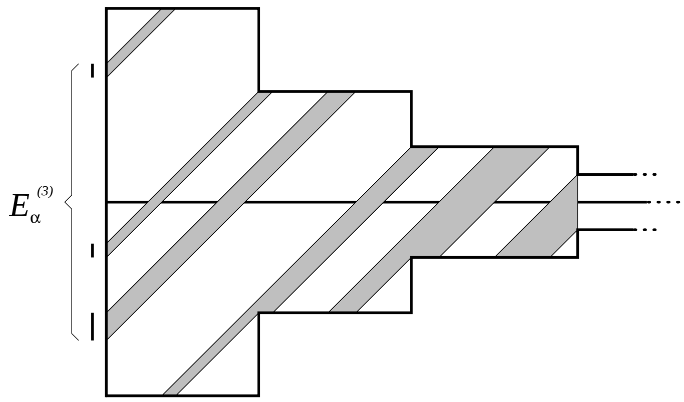

Figure 2.3: Construction of $E_\alpha^{(n)}$ as the backward evolution of the "aperture" G_n. The beam of orbits may split at singular vertices.

The subset of L on which $\mathcal{P}_\alpha$ is not defined is denoted by $E_\alpha := \bigcap_{n>0} E_\alpha^{(n)}$ and clearly $|E_\alpha| = 0$. Each point of this set is the limit of an infinite sequence of nested vanishing right-open intervals (the constituents of the sets $E_\alpha^{(n)}$). Elementary topology arguments allow us to assert an almost converse statement: each infinite sequence yields a point of E_α, unless the "pathological" property holds that the intervals eventually share their right extremes.

The orbits starting from such points will never collide with any vertical side of S_α (or P) and thus, as $t \to +\infty$, will go to infinity, maintaining a positive constant x-velocity. We call them *escape orbits*: their importance in this context will become especially clear in Chapter 3.

We have chosen this name for simplicity, but, strictly speaking, we consider only the asymptotic behavior of the forward semi-orbit. A glance at Fig. 2.1, however, shows at once that the backward semi-orbit having initial condition $(0, Y_0)$ is uniquely associated, by symmetry around the origin, to the forward semi-orbit of $(0, -Y_0)$.

REMARK. The above assertion needs to be better stated: although the manifold S_α is symmetric around the origin, the flow defined on it is not *exactly* invariant for time-reversal, as Fig. 2.1 seems to suggest. This is due to our convention in Section 2.1 about the continuity from above for the flow. The time-reversed motion on S_α is isomorphic to the motion on a manifold like S_α with the opposite convention (continuity from below). Nevertheless, little changes since only singular orbits (a null-measure set) are going to be affected by this slight asymmetry.

2.3 Preliminary results

Certain facts about the dynamics are immediate to derive. Let us start with the case of rational heights, $p_n \in \mathbb{Q}$. Many interesting infinite step billiards have this property, in particular our main example ($p_n = 2^{-n}$), treated in Chapter 4. One has

Proposition 2.1 *Fix $n \in \mathbb{N}$ and suppose $p_k \in \mathbb{Q}$, $\forall k \leq n$. Consider the billiard $P^{(n)}$. If $\alpha \in \mathbb{Q}$, all the trajectories are periodic. If $\alpha \notin \mathbb{Q}$, the flow is minimal and the Lebesgue measure is the unique invariant ergodic measure.*

PROOF. This proposition can be derived from quite a number of results in the literature (cf. Proposition 1.1). However, to give an exact reference, [G1, Thm. 3] contains the assertion, since $P^{(n)}$ is an *almost integrable* billiard table.

It may be interesting to remark that the ideas on which the proofs are based were already known sixty years ago, as [FK] witnesses. The invariant surface $S_\alpha^{(n)}$ is divided into a finite number of strips, that are either minimal sets or collections of periodic orbits (the two cases cannot occur simultaneously for an almost integrable billiard). These strips are delimited by *generalized diagonals*, that is, pieces of trajectory that connect two (possibly coincident) singular vertices of the invariant surface. The above is nowadays called *the structure theorem* for rational billiards, a sharp formulation of which is found, e.g., in [AG].

Using this, minimality is easily established when, for a given direction, no generalized diagonals and no periodic orbits are found. Q.E.D.

The primary consequence of Proposition 2.1 is

Proposition 2.2 *Let an infinite step billiard P with rational heights ($p_n \in \mathbb{Q}$, $\forall n$) be given. If $\alpha \in \mathbb{Q}$, a semi-orbit can be either periodic or unbounded. If $\alpha \notin \mathbb{Q}$, all semi-orbits are unbounded.*

PROOF. If $\alpha \in \mathbb{Q}$ and we had a non-periodic bounded trajectory, this would naturally correspond to a trajectory of $S_\alpha^{(n)}$, for some $n \in \mathbb{N}$, which has only periodic orbits. On the other hand, if $\alpha \notin \mathbb{Q}$, the dynamics over each $S_\alpha^{(n)}$ is minimal. Hence, every semi-trajectory reaches the abscissa $X = n$. Q.E.D.

One might wonder whether in the case of rational heights it is easy to find unbounded orbits for $\alpha \in \mathbb{Q}$. Escape orbits are natural candidates. For instance, in the model we treat in Chapter 4 we have at least one of them for every direction (Proposition 4.11). However, it might be interesting to construct less trivial examples, such as unbounded non-escape orbits. This be can done in the following way.

Consider a billiard table P whose heights $p_n \in \mathbb{Q}$ are yet to be determined. Now let a particle depart from $(0, 0)$, with slope $\alpha \in \mathbb{Q}$. Fix $p_1, \ldots, p_{n_1-1}$ in such a way that the particle hits no vertical walls before reaching the vertical line $x = n_1$; also choose n_1 in such a way that the particle approaches the line $x = n_1$ with a positive slope. Call (n_1, y_1) the intersection of the trajectory with this line. Now fix $p_{n_1} := y_1$. We have thus made our orbit singular. According to the conventions established at the beginning of this section, the particle will bounce against the vertex and reverse x-velocity. Now, by Proposition 2.2, two possible cases may occur: either the trajectory will reach again, after some time, the line $x = n_1$, and cross it, or it "will try" to become periodic in the truncated billiard $P^{(n_1)}$. In the latter case, in order to go back to $(0, 0)$, a non-singular vertex, the particle must hit another non-singular vertex, reverse velocity and run along the same trajectory in the opposite sense (Fig. 3.4 might help to clarify this point). In so doing, it will necessarily come back to the point (n_1, p_{n_1}), with a negative slope.

According to our conventions, this time it will pass through. In both cases, the particle has left $P^{(n_1)}$, having bounced at least once on a vertical wall. We can now repeat the same procedure to find some $n_2 > n_1$ and to fix $p_{n_1+1}, \ldots, p_{n_2-1}$, and then p_{n_2}. In the same way, after passing $x = n_1$, the particle will bounce at least once—at the singular vertex (n_2, p_{n_2})—and, after some time, move on to the right. This can be repeated indefinitely. What we end up having is a billiard with an unbounded non-escape orbit.

In the general case the previous proposition reads

Proposition 2.3 *In an infinite step billiard P, for almost all α's, all semi-orbits are unbounded, whereas periodic and unbounded trajectories can coexist for a zero-measure set of directions.*

Proof. This is just a generalization of the previous proof, although we are not dealing here with almost integrable tables.

For a given $\alpha \in \mathbb{R}^+$ and $n > 1$, call $\phi_{\alpha,t}^{(n)}$ the flow on $S_\alpha^{P^{(n)}}$, and let $\mathcal{M}_n := \{\alpha \mid \phi_{\alpha,t}^{(n)}$ is minimal$\}$. From Proposition 1.1,*(ii)* we get $|\mathcal{M}_n^c| = 0$. (The superscript ^{c} denotes here the complementary set w.r.t. to $\mathbb{R}^+$. In general, though, we will use the same symbol for the complement in other contexts as well.) Let $\mathcal{M}_\infty = \cap_{n>1} \mathcal{M}_n$. Clearly $|\mathcal{M}_\infty^c| = 0$. It is easy to see that, for all $\alpha \in \mathcal{M}_\infty$, every semi-orbit is unbounded.

As regards the last remark, it is not hard to think of a billiard which has at least a periodic and an escape orbit. (The skeptical reader can consider the exponential step billiard of Chapter 4, with $\alpha = 1$—for which it is trivial to see periodic orbits—and then apply Proposition 4.11.) Actually, playing with the arguments outlined before Proposition 2.3, one can easily produce step billiards that have both periodic and unbounded non-escape orbits. Q.E.D.

Chapter 3

Topological dynamics

It turns out that if the "infinite cusp" of a step billiard narrows down very quickly, there is a unique escape orbit for almost every direction. This is shown in Section 3.1. The behavior of this trajectory is studied in Section 3.2. This analysis is interesting in its own right and, more importantly, provides information about the topological behavior of all other orbits, which is presented in Section 3.3.

3.1 Uniqueness of the escape orbit

Theorem 3.1 *If the heights $\{p_n\}$ of an infinite step billiard P verify $p_{n+1} \leq \lambda p_n$, with*

$$0 < \lambda < \lambda_0 := \frac{\sqrt{6}-1}{5} \simeq 0.290\ldots,$$

then, for a.e. α, there exists a subsequence $\{n_j\}$ such that $n.i.(E_\alpha^{(n_j)}) = 1$.

The main corollary of this theorem is the following.

Corollary 3.2 *For almost all directions there is exactly one escape orbit.*

It might seem that Theorem 3.1 is just a technical lemma to ensure the validity of the above notable result. But this is not really so. It will be explained in Section 3.3 that Corollary 3.2 is not enough to prove the statements about the topological dynamics presented there. Furthermore, Lemma 4.10 in the next chapter will show that there are cases in which Corollary 3.2 holds but Theorem 3.1 does not.

For the rest of this chapter, all surfaces S_α^P will be identified with the set S^P of Fig. 2.1 on which we use the coordinates (X, Y). For instance, $\gamma_k(\alpha)$ will denote the semi-orbit on S^P_α starting at point the $(k,-p_k) = V_k$; this is the same as the semi-orbit on S^P starting at $(k,-p_k)$ with slope α.

We move on to the proofs of our claims.

PROOF OF THEOREM 3.1. We say that a trajectory reaches *directly* aperture G_n when this happens monotonically in the X-coordinate, that is, without intersecting any vertical walls.

Let $k < m$ be two natural numbers. We introduce the following sets:

$$\begin{aligned} A_{k,m} &:= \{\alpha \mid \gamma_k(\alpha) \text{ reaches directly } G_m\}; && (3.1)\\ B_m &:= \bigcup_{k=1}^{m-1} A_{k,m}. && (3.2)\end{aligned}$$

Therefore,

$$B_m^c = \{\alpha \mid \text{no } \gamma_k(\alpha),\ \text{with } 1 \le k \le m-1,\ \text{reaches directly } G_m\}. \qquad (3.3)$$

Also, let us define

$$\begin{aligned} C &:= \bigcap_{n\in\mathbb{N}} \bigcup_{m\ge n} B_m^c && (3.4)\\ &= \{\alpha \mid \exists\, \{n_j\} \ s.t.\ G_{n_j} \text{ is not reached by any } \gamma_k(\alpha), k \le n_j\}. \end{aligned}$$

When such a subsequence exists, the backward evolution of G_{n_j} does not split at any of the vertices V_k, $1 \le k \le n_j$. Hence $n.i.(E_\alpha^{(n_j)}) = 1$. Therefore establishing the theorem amounts to proving that $|C^c| = 0$. This is implied by the following:

$$\forall n \in \mathbb{N}, \quad \left| \bigcap_{m\ge n} B_m \right| = 0. \qquad (3.5)$$

In order to obtain (3.5), we introduce some notation, and a lemma. Given two sets A, I, with I bounded, denote $|A|_I := |A \cap I|/|I|$.

Lemma 3.3 *Assumptions and notations as in Theorem 3.1. There exists a* $\delta \in]0,1[$ *such that, for every (bounded) interval* $I \subset \mathbb{R}^+$, $\limsup_{m\to\infty} |B_m|_I \le \delta$.

The proof of this lemma will be given later. Now, proceeding by contradiction, let us suppose that $\left|\bigcap_{m\ge n} B_m\right| \ne 0$, for some n. By the Lebesgue's Density Theorem, almost all points of this set are points of density. Pick one: this means that there exists an interval I, around that point, such that

$$\left| \bigcap_{m \geq n} B_m \right|_I \geq \sigma > \delta. \tag{3.6}$$

Hence, $\forall m \geq n$, $|B_m|_I \geq \sigma$, which contradicts the lemma. This proves (3.5) and Theorem 3.1. Q.E.D.

Proof of Corollary 3.2. The previous theorem implies that, for a.a. α's, $\#E_\alpha = 0$ or 1. As already mentioned, the former case (no escape orbits), occurs if, and only if, the intervals $E_\alpha^{(n_j)}$ share their right extremes, for j large. This implies the existence of generalized diagonals (see Fig. 3.1). Excluding those cases only amounts to removing a null-measure set of directions. Q.E.D.

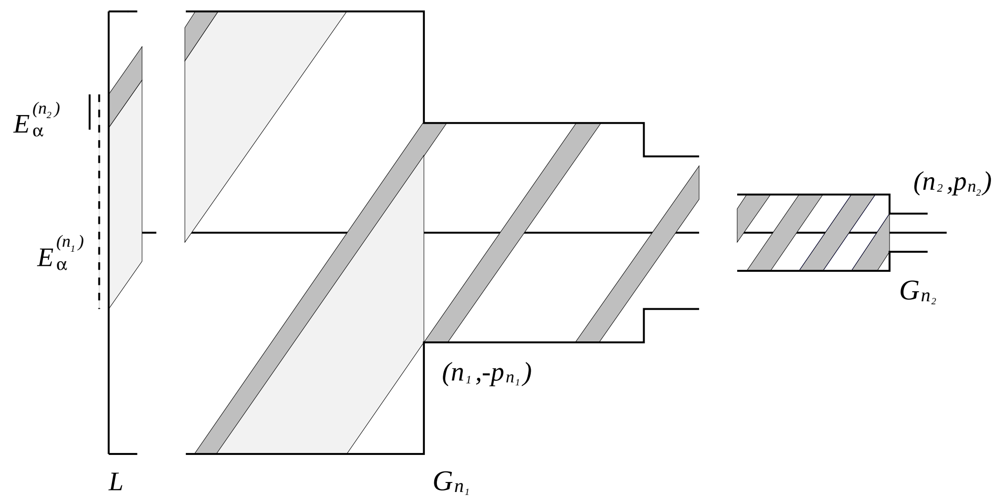

Figure 3.1: The fact that $E_\alpha^{(n_1)}$ and $E_\alpha^{(n_2)}$ have upper (equivalently right) extremes in common implies the existence of a generalized diagonal.

Proof of Lemma 3.3. In this proof we will heavily use the technique of billiard-unfolding; that is, in order to draw an orbit as a straight line in the plane, we reflect the billiard around one of its sides every time the orbit hits it, as shown in Fig. 3.2.

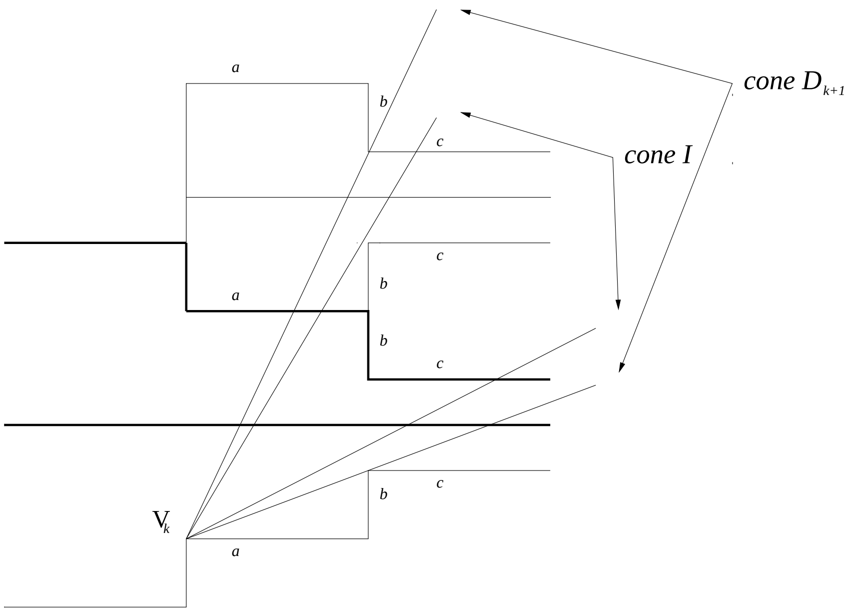

Figure 3.2: Unfolding of the billiard. Trajectories departing from a given singular vertex V_k are drawn as straight lines on the plane. Every time one of these hits a side of the billiard, a new copy of the billiard, reflected around that side, is drawn. Cones I and D_{k+1} are used in the proof of Lemma 3.3.

Fix k, and view I as a conical beam of trajectories departing from V_k (Fig. 3.2): this makes sense since these trajectories are in a one-to-one correspondence with their slopes. The goal is to exploit the geometry of the unfolded billiard to set up a recursive argument that will yield exponential bounds for $|A_{k,m}|_I$ (in m).

Here is how this works. In the unfolded-billiard plane, sketched in Fig. 3.3, let $l_k(\alpha)$ be the straight line of slope α passing through V_k. Take an $n > k$ and consider one copy of G_n, indicated as an "opening" in Fig. 3.3: $\widetilde{G}_n := \{n\} \times [r - p_n, r + p_n[$, for some $r \in \mathbb{R}^+$. The straight lines (departing from V_k) that cross $\widetilde{G}_n$ encounter a $2p_n$-periodic array of copies of G_{n+1}. We call them $\widetilde{G}^{(j)}_{n+1} := \{n+1\} \times [r - p_{n+1} + j \cdot 2p_n, r + p_{n+1} + j \cdot 2p_n[$; j assumes a finite number (say ℓ) of integer values. Define the interval

$$D_n := \{\alpha \mid l_k(\alpha) \cap \widetilde{G}_n \neq \emptyset\}. \tag{3.7}$$

Remark. It might be convenient to think of the elements of D_n as lines sharing the common point V_k: specifically the lines in D_n are those that cross $\widetilde{G}_n$. So it makes

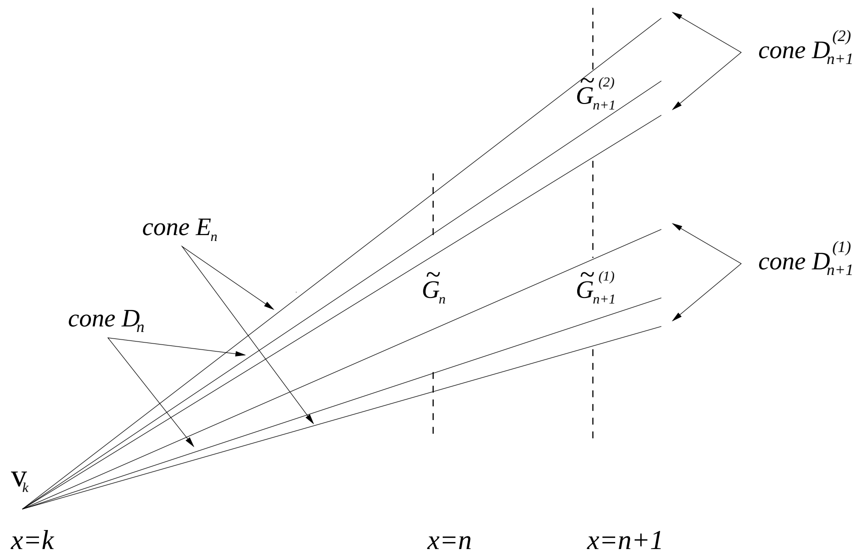

Figure 3.3: Proof of Lemma 3.3: Construction of the cones D_n and E_n in a roughly sketched unfolded-billiard plane.

sense to refer to D_n (and similar sets) as a cone. (Incidentally, $|D_n| = 2p_n/(n-k)$.) On the other hand, the one-to-one correspondence between $\alpha \in \mathbb{R}^+$ and $\gamma_k(\alpha)$, on one side, and $l_k(\alpha)$, on the other, should not lead one to think that the latter is the "lifting" of the former by means of the unfolding procedure. This is only the case when $\gamma_k(\alpha)$ reaches G_n. As a matter of fact, the inclusion $A_{k,n} \cap D_n \subseteq D_n$ is expected to be strict for general choices of D_n, $n > k+1$.

The cone D_n cuts a segment on the vertical line $x = n+1$. This segment includes some $\widetilde{G}^{(j)}_{n+1}$ and only intersects some other $\widetilde{G}^{(i)}_{n+1}$ (at most two, of course). Now expand D_n in such a way that the segment includes ℓ "full" copies of $G^{(j)}_{n+1}$; the resulting cone will be denoted by E_n. Fig. 3.3 shows that in this operation we might have to attach, on top and on bottom of D_n two cones of measure up to $2p_{n+1}/(n+1-k)$. Therefore:

$$\frac{|E_n|}{|D_n|} \le 1 + \frac{4p_{n+1}}{(n+1-k)|D_n|} \le \frac{2p_{n+1}+p_n}{p_n}. \tag{3.8}$$

We further define the set

$$D_{n+1} := \{\alpha \in E_n \mid l_k(\alpha) \cap \widetilde{G}^{(j)}_{n+1} \ne \emptyset, \text{ for some } j\} = \bigcup_{i=1}^{\ell} D^{(i)}_{n+1}, \tag{3.9}$$

where the $D^{(i)}_{n+1}$ are cones of measure $2p_{n+1}/(n+1-k)$. One has

$$\frac{|D_{n+1}|}{|E_n|} \le \frac{2 \cdot 2p_{n+1}}{2 \cdot 2p_{n+1} + 2(p_n - p_{n+1})} = \frac{2p_{n+1}}{p_{n+1} + p_n}. \tag{3.10}$$

In fact it is not hard to realize that the l.h.s. of (3.10) is largest when $\ell = 2$ and E_n is the juxtaposition of $D^{(1)}_{n+1}$, a beam of orbits that do not cross $x = n$, and $D^{(2)}_{n+1}$. This is the situation depicted in Fig. 3.3. In this case the central interval, $E_n \setminus D_{n+1}$, measures $2(p_n - p_{n+1})/(n+1-k)$, whence the second term of (3.10). Combining (3.8) and (3.10) we obtain

$$\frac{|D_{n+1}|}{|D_n|} \le \frac{2p_{n+1}(2p_{n+1} + p_n)}{p_n(p_{n+1} + p_n)} =: \beta_n. \tag{3.11}$$

At this point we notice that each $D^{(i)}_{n+1}$ as introduced in (3.9) is again a set of the type (3.7), with $n+1$ replacing n. Hence estimate (3.11) holds and $|D^{(i)}_{n+2}| \le \beta_{n+1}|D^{(i)}_{n+1}|$ with $D^{(i)}_{n+2}$ suitably defined as in the above construction. Defining $D_{n+2} := \cup_{i=1}^{\ell} D^{(i)}_{n+2}$, one has

$$\frac{|D_{n+2}|}{|D_{n+1}|} \le \frac{\sum_{i=1}^{\ell} |D^{(i)}_{n+2}|}{\ell \, |D^{(i)}_{n+1}|} \le \beta_{n+1}, \tag{3.12}$$

and the trick can continue.

We are now ready to implement the recursive argument: assume that some of the orbits of the cone I (based in V_k) cross G_{k+1}, that is, $A_{k,k+1} \cap I \ne \emptyset$ (if not, everything becomes trivial as we will see later). In the unfolded-billiard plane, enlarge $A_{k,k+1} \cap I$ until it fits the minimal number of copies of G_{k+1}, as in Fig. 3.2; call this new set D_{k+1}. This is by definition a finite union of intervals of the type (3.7), of fixed size. Hence inequality (3.11) applies and the definition/estimate algorithm can be carried on until we define, say, D_m. The only thing we need to know about this set is that $(A_{k,m} \cap I) \subseteq D_m$, which should be clear by construction (see also the previous remark). This fact and the repeated use of (3.11) yield

$$|A_{k,m} \cap I| \le |D_{k+1}| \prod_{i=k+1}^{m-1} \beta_i. \tag{3.13}$$

Let us consider our specific case: $p_{n+1} \le \lambda p_n$. From definition (3.11), one verifies that:

$$\beta_n = \frac{2(2(p_{n+1}/p_n) + 1)}{((p_n/p_{n+1}) + 1)} \le \frac{2(2\lambda + 1)}{\lambda^{-1} + 1} =: \beta. \tag{3.14}$$

It is going to be crucial later that β be less than 1. For λ positive, this amounts to $4\lambda^2 + \lambda - 1 < 0$, which is easily solved by

$$0 < \lambda < \lambda_1 := \frac{\sqrt{17}-1}{8} \simeq 0.390\ldots \tag{3.15}$$

Going back to the definition of D_{k+1}, and to (3.13), we see that it is possible to give an estimate of the measure of D_{k+1} in I, for large k. In fact, when $|I|$ is much bigger than $2p_k$, then it is clear (from Fig. 3.2, say) that I includes very many cones of size $2p_{k+1}$, placed on a $2p_k$-periodic array. As k grows, the density of these cones in I can be made arbitrarily close to $p_{k+1}/p_k \leq \lambda$. The precise statement then is: given any $\varepsilon > 0$, there exists a $q = q(\varepsilon) \in \mathbb{N}$ such that

$$\forall k \geq q, \quad |D_{k+1}| < (\lambda + \varepsilon)\, |I|. \tag{3.16}$$

Plugging this into (3.13), we obtain

$$\forall m > k \geq q \quad |A_{k,m}|_I \leq (\lambda + \varepsilon)\, \beta^{m-k-1}, \tag{3.17}$$

having used (3.14) as well. For the other values of k, we have no control over $C_k := |D_{k+1}|/|I|$ and we just write

$$\forall m, q > k \quad |A_{k,m}|_I \leq C_k\, \beta^{m-k-1}. \tag{3.18}$$

In the case $A_{k,k+1} \cap I = \emptyset$, which we did not consider before, (3.17)-(3.18) are trivial consequences of the fact that $A_{k,m} \cap I = \emptyset$, for all $m > k$.

We move on to the final estimation. Take $m > q$: from definition (3.2) we have, using (3.17) and (3.18),

$$\begin{aligned} |B_m|_I &\leq \sum_{k=1}^{q-1} |A_{k,m}|_I + \sum_{k=q}^{m-1} |A_{k,m}|_I \\ &\leq \sum_{k=1}^{q-1} C_k\, \beta^{m-k-1} + \sum_{k=q}^{m-1} (\lambda + \varepsilon)\, \beta^{m-k-1} \\ &\leq o(1) + \frac{\lambda + \varepsilon}{1 - \beta}, \end{aligned} \tag{3.19}$$

as $m \to \infty$. In the last inequality we have used twice the fact that $\beta < 1$. We impose the condition

$$\frac{\lambda}{1-\beta} = \frac{\lambda(\lambda+1)}{-4\lambda^2 - \lambda + 1} < 1. \tag{3.20}$$

For λ as in (3.15), the denominator is positive, hence (3.20) can be rewritten as $5\lambda^2 + 2\lambda - 1 < 0$, whose solutions are

$$0 < \lambda < \frac{\sqrt{6}-1}{5} = \lambda_0 < \lambda_1, \tag{3.21}$$

as in the statement of the lemma. For these values of λ, by virtue of (3.15), $1-\beta$ keeps away from 0. Therefore (3.20) implies that, for ε small enough, the last term in (3.19) can be taken less than a certain $\delta < 1$, whence the proof of Lemma 3.3. Q.E.D.

3.2 The backward part of the escape orbit

In this section, we explore the behavior of the escape orbits for $t \to -\infty$. This question happens to be crucial for the understanding of the dynamics on the billiard, as we shall see in the next section. All the claims we will make in the rest of Chapter 3 apply to billiards as in the statement of Theorem 3.1.

Let D_1 be the set of directions that satisfy Theorem 3.1 and Corollary 3.2, and let η_α denote the unique escape orbit, for $\alpha \in D_1$. As trivial as it is, we point out that the backward part of η_α cannot be periodic. Also, by Proposition 2.3, it cannot be bounded, for a.e. α: we call this set D_2. What about the possibility for η_α to escape to ∞ in the past, with a constant negative x-velocity?

Lemma 3.4 *For a.a. α, η_α does not intersect any vertex.*

Proof. Consider a vertex V of S and let $\gamma_V(\alpha)$ be its forward semi-orbit. For a fixed finite sequence of sides, $\Lambda := (\Lambda_1, \Lambda_2, \ldots, \Lambda_\ell)$, we define

$$A_{V,m}(\Lambda) := \{\alpha \mid \gamma_V(\alpha) \text{ hits } \Lambda_1, \ldots, \Lambda_\ell \text{ and then reaches directly } G_m\}. \tag{3.22}$$

This means that these trajectories do not hit any vertical wall after leaving Λ_ℓ and before reaching G_m. Notice the similarities with definition (3.1). As a matter of fact, if $V = V_k$ for some k, then $A_{V_k,m}(\emptyset) = A_{k,m}$. However, for most sequences Λ, (3.22) defines the empty set. For example, Λ can be *incompatible* in the sense that no orbit can go from Λ_i to Λ_{i+1} without crossing other sides in the meantime. But, even for compatible sequences, if G_m does not lie to the right of Λ_ℓ, obviously $A_{V,m}(\Lambda) = \emptyset$. To avoid this latter case, we fix $m_o = m_o(V, \Lambda)$ bigger than the largest Y-coordinate in Λ_ℓ. For $m \geq m_o$, $A_{V,m+1}(\Lambda) \subset A_{V,m}(\Lambda)$. Let us then define

$$\begin{aligned} B_V(\Lambda) &:= \{\alpha \,|\, \gamma_V(\alpha) \text{ hits } \Lambda_1, \ldots, \Lambda_\ell \text{ and then escapes to } \infty\} \\ &:= \bigcap_{m=m_o}^{\infty} A_{V,m}(\Lambda), \end{aligned} \tag{3.23}$$

Working in the unfolded-billiard plane and identifying directions with orbits, it is not hard to realize that $A_{V,m_o}(\Lambda)$ is made up of a finite number of intervals/cones, each of which reaches a copy of G_{m_o} after hitting a certain sequence of sides ($\Lambda_1, \ldots, \Lambda_\ell$, $\Lambda_{\ell+1}, \ldots, \Lambda_n$) (the first ℓ sides are common to all cones and the others can only be horizontal). It is possible that some of these beams of trajectories intersect the corresponding copy of G_{m_o} only in a proper sub-segment. Let us fix this situation by enlarging any such beam until it covers the whole segment. We call $D \supseteq A_{V,m_o}(\Lambda)$ the union of these new cones.

Proceeding exactly as in the proof of Lemma 3.3 (see in particular (3.13) and (3.18)) we get, for $m > m_o$,

$$|A_{V,m}(\Lambda)| \leq |D| \, \beta^{m-m_o},$$

with $\beta < 1$. Hence, for $m \to \infty$, $|A_{V,m}(\Lambda)| \to 0$. By (3.23), $|B_V(\Lambda)| = 0$, and the set

$$\bigcup_{\substack{V \\ \text{vertex}}} \; \bigcup_{\substack{\Lambda \text{ finite} \\ \text{sequence}}} B_V(\Lambda)$$

has measure zero. This is the complement of set of directions as in the statement of Lemma 3.3. Q.E.D.

Let D_3 denote the set mentioned above. With a certain lack of originality, we call *typical* any $\alpha \in D := D_1 \cap D_2 \cap D_3$, that is, any direction that has all the properties we have analyzed so far.

Corollary 3.5 *For a.a. α's, the billiard flow $\phi_{\alpha,t}$ around the escape orbit is a local isometry. This means that, fixed a $z_0 \in \eta_\alpha$, then $\forall T > 0$, $\exists \varepsilon > 0$ s.t.*

$$|z - z_0| \leq \varepsilon \Longrightarrow |\phi_{\alpha,t}(z) - \phi_{\alpha,t}(z_0)| = |z - z_0| \quad \forall t \in [-T/2, T/2].$$

PROOF. Easy consequence of Lemma 3.4, since $\phi_{\alpha,t}$ is isometric far from the singular vertices, for $\alpha \in D_1 \cap D_3$. Q.E.D.

We borrow some notation from [L] and call *oscillating* all unbounded non-escape (semi-)orbits.

Theorem 3.6 *For a typical α, the unique escape orbit is oscillating in the past.*

PROOF OF THEOREM 3.6. Since $\alpha \in D_3$, η_α is non-singular, by Lemma 3.4. Thus the symmetry arguments outlined at the end of Section 2.2 apply. Moreover η_α is unbounded in the past, because $\alpha \in D_2$. Suppose now that η_α^-, some past semi-orbit, escapes: this corresponds, by reflection, to a forward escape semi-orbit. Then the uniqueness hypothesis shows that the reflected image of η_α^- must coincide with some η_α^+. In other words, η_α is symmetric around the origin in S_α^P, which means that in P it is run over twice, once for each direction. The situation is illustrated, for both P and S_α^P, in Fig. 3.4. One gets easily convinced that the only way to realize this case is that the trajectory has a point in which the velocity is inverted. This can only be a non-singular vertex. But $\alpha \in D_3$ and Lemma 3.4 claims that this is impossible. Q.E.D.

3.3 The escape orbit as an "attractor"

In the remainder we fix a direction $\alpha \in D$ and, for simplicity, we drop the subscript α and the superscript P from all the notation. Remember that we are considering a billiard P verifying the hypothesis of Theorem 3.1.

We show here how it is possible to use the statements of Section 3.2 on the special orbit η to derive a certain amount of information about the topology of all the other trajectories.

On η we fix the *standard initial condition* $z_0 = (0, Y_0)$ as the last intersection point with L, before the orbit escapes towards ∞.

The crucial fact, as will be noted, is Theorem 3.1, which roughly states that not only is there just one initial point that takes a trajectory to infinity, but also there

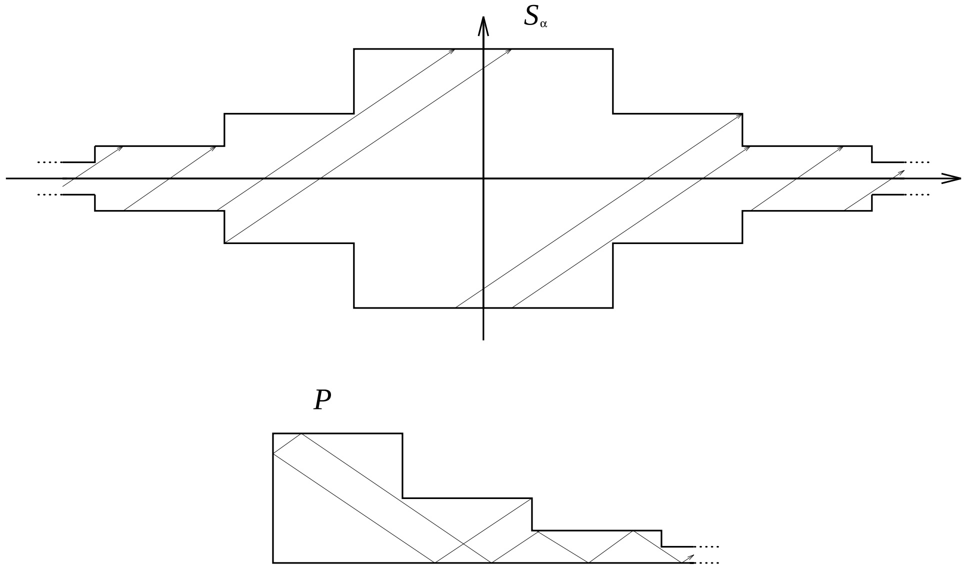

Figure 3.4: A trajectory can be run over twice only if it contains a non-singular vertex. For a.a. α's this is the only possibility to have an orbit which escapes both in the past and in the future.

is just one neighborhood—necessarily around that point—that takes a trajectory far enough. This is the idea behind the next result.

Lemma 3.7 *Let $\alpha \in D$. Fixed an orbit γ and two numbers $\varepsilon, T > 0$, there exists a $w \in \gamma \cap L$, such that*

$$|\phi_{\alpha,t}(w) - \phi_{\alpha,t}(z_0)| = |w - z_0| \leq \varepsilon \quad \forall t \in [-T/2, T/2],$$

where z_0 is the standard initial condition on η.

Furthermore, call $\widetilde{\eta}$ the image of η through a symmetry around the origin. If $\gamma \neq \eta, \widetilde{\eta}$, w can be chosen arbitrarily far in the past or in the future of γ. For $\gamma = \eta$, resp. $\widetilde{\eta}$, w can be chosen arbitrarily far in the past, resp. future.

Proof. Since α is typical, we can apply Corollary 3.5 with z_0 fixed as above. This will return an ε' (depending on T) such that all points as close to z_0 as ε' remain such under the flow, within a time T. Assume $\varepsilon' \leq \varepsilon$ (if not, $\varepsilon' := \varepsilon$ will do). We need to find a point of γ in the interval $[Y_0 - \varepsilon', Y_0 + \varepsilon'] \subseteq L$. Recalling Theorem 3.1, consider the subsequence $\{G_{n_j}\}$ of apertures whose backward beam of trajectories does not split at any vertex before reaching L. Take a j such that $2p_{n_j} \leq \varepsilon'$. Since γ is unbounded, we can find a point $u \in \gamma \cap G_{n_j}$. Call w the last intersection point of γ with L, before u is reached. From the non-splitting property of G_{n_j}, $|w - z_0| \leq \varepsilon'$. Corollary 3.5 shows that this is the sought w.

If $\gamma \neq \eta, \widetilde{\eta}$, Proposition 2.3 states that *each* semi-trajectory of γ is oscillating (for $\alpha \in D_2$): therefore u (and so w) can be chosen with as much freedom as claimed in the last statement of the lemma. As for η, only the backward part oscillates, whereas $\widetilde{\eta}$ oscillates in the future, being the escape orbit in the past. Q.E.D.

REMARK. We stress once again that the above is more than an easy corollary of Proposition 2.3: not only do γ and η get close near infinity, being both squeezed inside the narrow "cusp", but, to be so, they must have been as close for a long time.

A number of trivially checkable consequences of Lemma 3.7 are listed in the sequel. Recall the definitions of ω-limit and α-limit of an orbit as the sets of its accumulation points in the future and in the past, respectively [W, Def. 5.4].

Corollary 3.8 *With the same assumptions and notation as above,*

(i) The escape orbit η is contained in the ω-limit and in the α-limit of every orbit, other than η and $\widetilde{\eta}$.

(ii) The escape orbit η is contained in its own α-limit and in the ω-limit of $\widetilde{\eta}$.

(iii) Every invariant continuous function is constant.

(iv) The flow is minimal if, and only if, the escape orbit is dense.

Of course, one would like to prove one definite topological property of the flow ϕ_t, such as minimality or at least topological transitivity. Our techniques do not seem to accomplish this. However, they do furnish a picture of how chaotic the motion on the billiard can be. In fact, the "attractor" that η has been proven to be is certainly far from simple. Either it densely fills the whole invariant surface S_α, or it is a fractal set.

Theorem 3.9 *For a typical direction consider the corresponding flow on* S_α*. Denote* $L_\eta := \eta \cap L$*, the "trace" of the escape orbit on the usual Poincaré section. Then its closure in* L *(denoted by* $\overline{L_\eta}$*) is either the entire* L *or a Cantor set.*

Proof of Theorem 3.9. Assume $\overline{L_\eta} \neq L$. This set is closed. We are going to show it also has empty interior and no isolated points, that is, it is Cantor. In the remainder, by interval we will always mean a segment of L.

Suppose the interior of our set is not empty. Then there exists an open interval $I \subseteq \overline{L_\eta}$ containing a point z of η. Now, in the complementary set of $\overline{L_\eta}$, select a point w whose orbit is non-singular. Let w evolve, e.g., in the future. By Corollary 3.8,*(i)* applied to z, there is a $t > 0$ such that $\phi_t(w) \in I$. By the choice of w, we can find an open interval J, such that $w \in J$, $J \cap \overline{L_\eta} = \emptyset$ and so small that ϕ_t maps J isometrically into I. This implies that $J \subset \overline{L_\eta}$, which is a contradiction.

To show that there are no isolated points: if $z \in \overline{L_\eta} \setminus L_\eta$, there is nothing to prove; if $z \in L_\eta$, then Corollary 3.8,*(ii)* will do. Q.E.D.

Chapter 4

The exponential step billiard

Infinite step billiards were first introduced in [DDL1]. Treated in detail there was the case $p_n = 2^{-n}$ which was christened the exponential step billiard.

We report the outcomes of that work in this chapter to show how the topological properties we have analyzed in Chapter 3 work in a specific case. Actually, the exponential billiard does not satisfy the assumptions of Section 3.1. Nevertheless the same kinds of results as in Sections 3.1-3.2 can be shown, due to the evident scaling symmetry that the exponential table possesses. This symmetry enables one to use rather elementary techniques to complete a thorough analysis of the escape orbits for the model at hand. Section 4.1 covers this.

Understandably, the exponential billiard is just one of the very many infinite step tables that validate the claims of Section 3.3 without verifying the hypothesis of Theorem 3.1.

Let us start with the following construction: suppose that a trajectory γ on S_α starts from L and reaches directly the opening $G_n = \{n\} \times [-2^{-n}, 2^{-n}[$ (cf. proof of Theorem 3.1). Let us denote with $Y_n \in [-2^{-n}, 2^{-n}[$ the ordinate of the point at which γ crosses G_n. Within the box $]n, n+1[\times[-2^{-n}, 2^{-n}[$, the motion is a simple translation. Hence

$$Y_{n+1} = Y_n + \alpha \pmod{2^{-n+1}}, \tag{4.1}$$

with (mod r) meaning the *unique point* in $[-r/2, r/2[$ representing the class of equivalence in $\mathbb{R}/r\mathbb{Z}$, rather than the class of equivalence itself. The trajectory γ will cross G_{n+1} if, and only if,

$$Y_{n+1} \in [-2^{-(n+1)}, 2^{-(n+1)}[. \tag{4.2}$$

Setting $y_n := 2^{n-1}Y_n$, relation (4.1) becomes

$$y_{n+1} = 2y_n + 2^n\alpha \pmod 2, \tag{4.3}$$

and the trajectory will cross G_{n+1} if, and only if,

$$y_{n+1} \in \left[-\frac{1}{2}, \frac{1}{2}\right[. \tag{4.4}$$

The recursion relation (4.3) can be easily proven by induction to yield

$$y_{n+1} = 2^{n+1}y_0 + (n+1)2^n\alpha \pmod 2, \tag{4.5}$$

where the numbers $y_k \in [-1/2, 1/2[$ now represent the (rescaled) intersections of the trajectory with the vertical openings G_k.

The transformation $T_{n,\alpha} : [-1/2, 1/2[\longrightarrow [-1, 1[$,

$$T_{n,\alpha}(y) := 2y + 2^n\alpha \pmod 2 \tag{4.6}$$

is called the *rescaled transfer map.*

4.1 Escape orbits for the exponential billiard

The definition of $T_{n,\alpha}$ shows that, due to the symmetry of our table, what really matters (as far as the escape orbits are concerned) is not the "constant of motion" α, but rather $\alpha \pmod 2$, the difference resulting only in the orbits winding more times around the boxes $]n, n+1[\times[-2^{-n}, 2^{-n}[$. As a matter of fact, we will see later that virtually all the information we need is stored in the binary expansion of $\alpha \pmod 2$.

We start out with some easy statements, using the rescaled coordinates y_k, unless otherwise specified.

Lemma 4.1 *If $\alpha = 2k$, $k \in \mathbb{Z}$, only one escape orbit exists and its initial condition is $y_0 = 0$.*

PROOF. $T_{n,2k} = T_{n,0}$ $\forall n \in \mathbb{N}$. The sequence $\{y_n\}$, in this case, is given by $y_n = 2^n y_0 \pmod 2$. If $y_0 = 0$, all y_n are null and the corresponding trajectory escapes, according to (4.4). If $y_0 \neq 0$, there exists a k such that $|y_k| > 1/2$. Q.E.D.

Corollary 4.2 *If $\alpha = k\,2^{-j}$, with k odd, j a non-negative integer, only one escape orbit occurs. This orbit intercepts V_j.*

PROOF. The portion of the manifold S_α to the right of the $(j+1)$-th aperture looks like S_α itself, modulo a scale factor equal to $2^{-(j+1)}$. Furthermore, in that region, and subject to the above rescaling, the transfer map is equivalent to the one we have seen in the previous lemma. In fact, for $n \geq j+1$, $T_{n,\alpha} = T_{n,0}$. So, to the part of the escape orbit after G_{j+1}, we can apply that lemma and conclude that the escape trajectory is unique and $y_{j+1} = 0$ holds. Now, we know that α is indeed equal to $(2k'+1)2^{-j}$. Inverting (4.3) with $y_{j+1} = 0$, we get $y_j = -k' - 1/2 - p$, for some integer p. By (4.4) $y_j \in [-1/2, 1/2[$. Hence $y_j = -1/2$, which proves the second part of the lemma. Q.E.D.

In Lemma 4.1 we have encountered the case in which $\{T_{n,\alpha}\}$ is a sequence of identical maps. Considering the more general case of a periodic sequence of maps will yield a useful tool to detect the presence of more than one escape orbit.

Observe that, when $\alpha = 2k/(2^m - 1)$ with k, m positive integers, one gets $2^m\alpha = \alpha \pmod 2$. In this case we have a periodic sequence of transfer maps of period m, that is, $T_{pm,\alpha} = T_{0,\alpha}$ for all integers $p > 0$. For such directions, then, one method for finding escape orbits may be the following: Set $M_{m,\alpha} := T_{m-1,\alpha} \circ \cdots \circ T_{0,\alpha}$. As in (4.5) it turns out that $M_{m,\alpha}(y) = 2^m y + m2^{m-1}\alpha \pmod 2$. Let us now find the fixed points of this map. Consider a trajectory having one of these points as initial datum. If it crosses all openings between G_1 and G_m, then the sequence of crossing points $y_0, \ldots, y_{m-1}$ will be indefinitely repeated and the trajectory will escape.

Let us apply this technique to the case $k = 1$ and $m = 2$, that is, $\alpha = 2/3$. The fixed points of the map $M_{2,2/3}$ are the points $y \in [-1, 1[$ such that $y = 4y + 8/3 + 2p$, $p \in \mathbb{Z}$. They are

$$y^{(0)} = -\frac{8}{9}; \quad y^{(1)} = -\frac{2}{9}; \quad y^{(2)} = \frac{4}{9}. \tag{4.7}$$

Since $|y^{(0)}| = |M_{2,2/3}\, y^{(0)}| > 1/2$, that solution has to be discarded. Instead, $y^{(1)} =: y_0^{(1)}$ is accepted since

$$y_1^{(1)} \quad = \quad 2y_0^{(1)} + \frac{2}{3} \pmod 2 = \frac{2}{9} \in \left[-\frac{1}{2}, \frac{1}{2}\right[\tag{4.8}$$

$$y_2^{(1)} \quad = \quad y_0^{(1)} \in \left[-\frac{1}{2}, \frac{1}{2} \right[. \tag{4.9}$$

It turns out that the same holds for $y^{(2)}$.

Thus, for $\alpha = 2/3$ there are at least two escape orbits whose initial conditions in the non-rescaled coordinates are $Y_0 = -4/9$ and $Y_0 = 8/9$. As orbits of S_α they are distinct, but is this still true if we consider the corresponding orbits in the billiard P?

Lemma 4.3 *For $\alpha = 2/3$, the two escape orbits with initial conditions $Y_0 = -4/9$ and $Y_0 = 8/9$ have distinct projections on P.*

PROOF. Suppose that the two escape orbits coincide in P. Then the backward part of the orbit that starts at $Y = -4/9$ must get to $Y = 8/9$, after several oscillations. That is, one trajectory turns into the other, by time reversal. According to the fact that a backward semi-orbit having initial condition $(0, Y_0)$ is associated to the forward semi-orbit of $(0, -Y_0)$ (see Section 2.2), the geometry of S_α implies that

$$\frac{8}{9} = -\frac{4}{9} + \sum_{i=1}^{j} \left(\frac{2}{3} - \frac{m_i}{2^{q_i}} \right), \tag{4.10}$$

where m_i, q_i are non-negative integers and j is the number of rectangular boxes visited by the backward semi-orbit before reaching the point with coordinate $Y = 8/9$. (Recall that in each box the variation of the Y-coordinate is $\alpha \pmod{2^{-q_i}}$.) Rearranging this formula we obtain

$$\frac{4}{9} = \frac{2}{3} j - \frac{m}{2^q}, \tag{4.11}$$

for some non-negative integers m and q. For any choice of m, q and j, the two sides are distinct. This contradicts our initial assumption. Q.E.D.

We will see later (Corollary 4.9) that there are no more than two escape orbits in S_α, for any α. We summarize everything about the case $\alpha = 2/3$ in the following assertion.

Proposition 4.4 *Along the direction $\alpha = 2/3$ there are two distinct escape orbits.*

We now turn to the study of the generic case. In view of the arguments of Section 2.2 (see in particular Fig. 2.3), and especially recalling the proof of Theorem 3.1, we naturally resolve to analyze the trajectories γ_p, i.e., the forward semi-orbits starting from the singular vertices $V_p = (p, -2^{-p})$.

First of all, we consider those α's for which γ_0 reaches directly G_n: We look at Fig. 4.1, which displays the unfolding of P. A direct evaluation with a ruler furnishes the answer, that runs as follows:

$$n = 1) \quad \frac{1}{2} \leq \alpha \pmod 2 < \frac{3}{2} \tag{4.12}$$

$$n \geq 2) \quad \begin{cases} \dfrac{1}{2} \leq \alpha \pmod 2 < \dfrac{1}{2} + \dfrac{1}{n2^n}; \\[2ex] 1 - \dfrac{1}{n2^n} \leq \alpha \pmod 2 < 1 + \dfrac{1}{n2^n}; \\[2ex] \dfrac{3}{2} - \dfrac{1}{n2^n} \leq \alpha \pmod 2 < \dfrac{3}{2}. \end{cases} \tag{4.13}$$

Due to the self-similarity of our table, we can write down the analogous inequalities for every other vertex V_p, $p \geq 1$, by rescaling (4.12) and (4.13). Thus γ_p crosses G_m, $m > p$, if, and only if,

$$m = p+1) \quad \frac{1}{2^{p+1}} \leq \alpha \pmod{2^{-p+1}} < \frac{3}{2^{p+1}} \tag{4.14}$$

$$m \geq p+2) \quad \begin{cases} \dfrac{1}{2^{p+1}} \leq \alpha \pmod{2^{-p+1}} < \dfrac{1}{2^{p+1}} + \dfrac{1}{(m-p)2^m}; \\[2ex] \dfrac{1}{2^p} - \dfrac{1}{(m-p)2^m} \leq \alpha \pmod{2^{-p+1}} < \dfrac{1}{2^p} + \dfrac{1}{(m-p)2^m}; \\[2ex] \dfrac{3}{2^{p+1}} - \dfrac{1}{(m-p)2^m} \leq \alpha \pmod{2^{-p+1}} < \dfrac{3}{2^{p+1}}. \end{cases} \tag{4.15}$$

Working out these relations is essentially all we need to do for the rest of this section. From now on, when we say that γ_p reaches or crosses aperture G_m, we will always mean *directly.*

Lemma 4.5 *If γ_p is an escape orbit then it is the only escape orbit.*

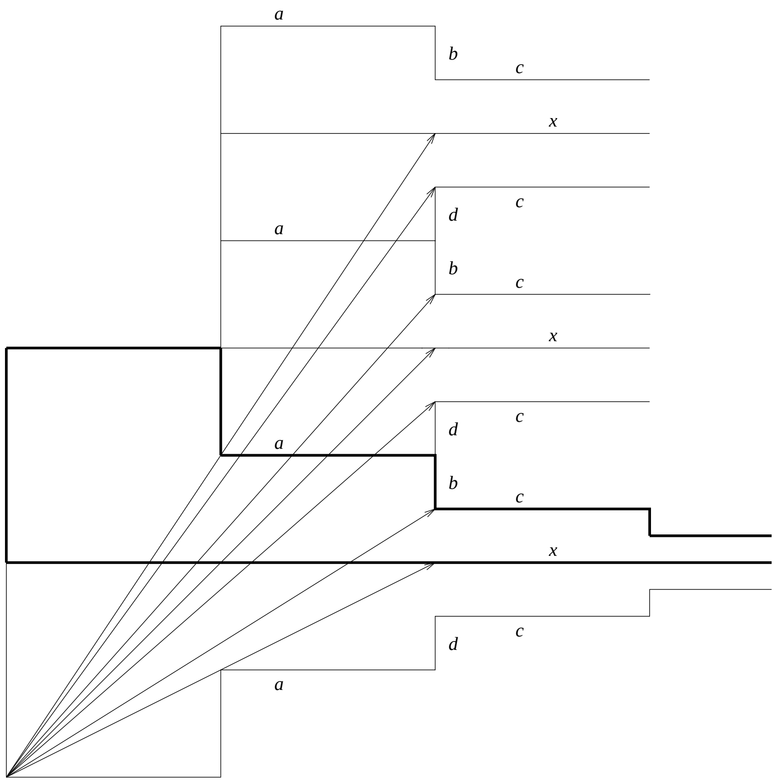

Figure 4.1: Range of directions for which the orbit starting from the leftmost bottom vertex reaches directly aperture G_n (case $n = 2$ is displayed). The exponential step billiard has been unfolded on the plane.

Proof of Lemma 4.5. It follows from (4.14)-(4.15) that γ_p is an escape orbit if, and only if, $\alpha \in \{2^{-p}, 2^{-p-1}\} \pmod{2^{-p+1}}$.

In fact, fix for simplicity $p = 0$, and consider Fig. 4.1 and (4.13): restricting to $\alpha \pmod 2$, three cones occur, for $n \geq 2$. Fixing one cone relative to a copy of G_n, the geometry of the billiard implies that one, and only one, sub-cone will also reach a copy of G_{n+1}. Eventually, for $n \to +\infty$, these three cones narrow down to the values $\alpha = 1/2, 1, 3/2$, the last of which is rejected for our convention on the continuation of singular orbits.

Now, for $\alpha \in \{2^{-p}, 2^{-p-1}\} \pmod{2^{-p+1}}$, Corollary 4.2 states that there is only one escape orbit. Q.E.D.

Lemma 4.6 *Let m, p be two non-negative integers with $m \geq p + 2$. If γ_p crosses G_m,* *then either γ_{p+1} does not reach G_{p+2} or it crosses G_m, as well.*

PROOF. It suffices to prove the statement for $p = 0$ and the reader can easily get convinced that the actual result follows by a rescaling. Set

$$\begin{aligned} I_m^{(1)} &:= \bigcup_{j\in\mathbb{N}} \left[\frac{1}{2}+2j, \frac{1}{2}+\frac{1}{m2^m}+2j\right[, \\ I_m^{(2)} &:= \bigcup_{j\in\mathbb{N}} \left[1-\frac{1}{m2^m}+2j, 1+\frac{1}{m2^m}+2j\right[, \\ I_m^{(3)} &:= \bigcup_{j\in\mathbb{N}} \left[\frac{3}{2}-\frac{1}{m2^m}+2j, \frac{3}{2}+2j\right[. \end{aligned} \tag{4.16}$$

From (4.13), γ_0 crosses G_m if, and only if, $\alpha \in I_m := I_m^{(1)} \cup I_m^{(2)} \cup I_m^{(3)}$. If $\alpha \in I_m^{(2)}$ then γ_1 does not cross G_2. In fact, (4.14) states that γ_1 crosses G_2 if, and only if, $\alpha \in B := \bigcup_{k\in\mathbb{N}}[1/4+k, 3/4+k[$. So we have to prove that the sets $I_m^{(2)}$ and B have empty intersection. This is the case, because $I_m^{(2)}$ is made up of intervals of center $2j+1$ and radius $1/(m2^m)$, and B is made up of intervals of center $1/2+k$ and radius $1/4$, so that

$$\mathrm{dist}(I_m^{(2)}, B) \geq \frac{1}{2} - \left(\frac{1}{m2^m}+\frac{1}{4}\right) > 0 \quad \text{for } m \geq 2. \tag{4.17}$$

It remains to analyze the case $\alpha \in C := I_m^{(1)} \cup I_m^{(3)}$. Relations (4.15) tell us that γ_1 crosses G_m if, and only if,

$$\begin{aligned} \alpha \in D := \bigcup_{k\in\mathbb{N}} \Bigg(& \left[\frac{1}{4}+k, \frac{1}{4}+\frac{1}{(m-1)2^m}+k\right[\\ & \cup \left[\frac{1}{2}-\frac{1}{(m-1)2^m}+k, \frac{1}{2}+\frac{1}{(m-1)2^m}+k\right[\\ & \cup \left[\frac{3}{4}-\frac{1}{(m-1)2^m}+k, \frac{3}{4}+k\right[\Bigg). \end{aligned} \tag{4.18}$$

We have to prove that $C \subseteq D$. We can visualize the sets C and D as periodic structures on the line whose fundamental patterns have, respectively, lengths 2 and 1 (with common endpoints). Therefore, defining $\widehat{C} := C \cap [0,2] = [1/2, 1/2+1/(m2^m)[\cup [3/2 - 1/(m2^m), 3/2[$ and $\widehat{D} := D \cap [0,2]$, all we have to do is to show that $\widehat{C} \subseteq \widehat{D}$. Deducing the shape of $\widehat{D}$ from (4.18), the result follows from the trivial relations:

$$\left[\frac{1}{2}, \frac{1}{2}+\frac{1}{m2^m}\right[\subset \left[\frac{1}{2}-\frac{1}{(m-1)2^m}, \frac{1}{2}+\frac{1}{(m-1)2^m}\right[, \tag{4.19}$$

$$\left[\frac{3}{2}-\frac{1}{m2^m}, \frac{3}{2}\right[\subset \left[\frac{3}{2}-\frac{1}{(m-1)2^m}, \frac{3}{2}+\frac{1}{(m-1)2^m}\right[. \tag{4.20}$$

Q.E.D.

Lemma 4.7 *Again $m \geq p+2$. If γ_p crosses G_m, then for all $p+2 \leq n \leq m$, γ_n does not reach G_{n+1}.*

PROOF. As before, we give the proof only for the case $p = 0$. The orbit γ_0 crosses G_m if, and only if, $\alpha \in I_m$, defined in the proof of the previous lemma, whereas γ_n crosses G_{n+1} if, and only if,

$$\alpha \in J_n := \bigcup_{k \in \mathbb{N}} \left[\frac{1}{2^{n+1}} + \frac{k}{2^{n-1}}, \frac{3}{2^{n+1}} + \frac{k}{2^{n-1}} \right[. \tag{4.21}$$

If I_m and J_n have empty intersection, for all $2 \leq n \leq m$, then the lemma is proven. Proceeding as in the first part of Lemma 4.6, we see that I_m is strictly contained in the union of all intervals of center $q/2$ and radius $1/(m2^m)$, while the intervals constituting J_n have center $2^{-n} + k2^{-n+1}$ and radius 2^{-n-1}. Thus, for $3 \leq n \leq m$,

$$\mathrm{dist}(I_m, J_n) \geq \frac{1}{2^n} - \left(\frac{1}{m2^m} + \frac{1}{2^{n+1}} \right) > 0. \tag{4.22}$$

If $n = 2$, (4.22) becomes an equality, but the fact that our intervals are right-open ensures nevertheless that $I_m \cap J_2 = \emptyset$. Q.E.D.

One way to memorize the previous technical lemmas may be as follows. The fact that γ_p crosses G_m influences all γ_n's, for n between $p+1$ and m: if γ_{p+1} wants to "take off" (that is, reach some apertures), then it is forced to follow, and possibly pass, γ_p; while the γ_n's with $n \geq p+2$ cannot even take off.

We now enter the core of the arguments: recall the notation $n.i.$ to designate the number of disjoint intervals that constitute a set.

Lemma 4.8 *Fix $\alpha > 0$. Either there exists an integer q such that $n.i.(E_\alpha^{(n)}) = 2$ for all $n \geq q$, or there is a sequence $\{n_j\}$ such that $n.i.(E_\alpha^{(n_j)}) = 1$.*

PROOF. The set of α's with the property that $n.i.(E_\alpha^{(n)}) = 2$ for $n \geq q$ is not empty. In fact, by direct computation, it is easy to verify that for $\alpha = 2/3$ every γ_n crosses G_{n+1} but not G_{n+2}, so that $n.i.(E_\alpha^{(n)}) = 2$ for all $n > 0$.

Now, suppose there exists a sequence $\{m_j\}$ with $n.i.(E_\alpha^{(m_j)}) \neq 2$. We can assume $n.i.(E_\alpha^{(m_j)}) \geq 3$, otherwise, maybe passing to a subsequence, we would have $n.i.(E_\alpha^{(m_j)}) = 1$ and we would be done. If we fix an m_j, there are at least two singular orbits that cross G_{m_j}. Let $0 < p_j \leq m_j - 2$ be the smallest integer such that γ_p crosses G_{m_j}. It follows from Lemma 4.7 that only γ_{p_j} and γ_{p_j+1} cross G_{m_j}. Therefore $n.i.(E_\alpha^{(m_j)}) = 3$.

At this point we have three cases: γ_{p_j} and γ_{p_j+1} are both escape orbits; one of them escapes and the other is reflected; they are both reflected.

In the first case Lemma 4.5 ensures that γ_{p_j} and γ_{p_j+1} coincide and Lemma 4.7 implies that no γ_n with $n > p_j + 1$ can "take off". Hence $n.i.(E_\alpha^{(m)}) = 2$ for all $n \geq p_j + 1$, contradicting our assumption. The second case is hardly different: call G_{n_j} the first aperture that γ_{p_j} cannot reach (in fact Lemma 4.6 implies that, of the two, γ_{p_j+1} must be the escaping trajectory). Therefore, using again Lemma 4.7, $n.i.(E_\alpha^{(m)}) = 2$ for all $n \geq n_j$, a contradiction as before. In the last case, call G_{n_j} the first aperture which is not reached by γ_{p_j+1}, and thus not even by γ_{p_j}. (Lemma 4.6 claims that γ_{p_j+1} goes farther than γ_{p_j}.) Another application of Lemma 4.7 proves that $n.i.(E_\alpha^{(n_j)}) = 1$. Proceeding inductively we find a sequence of integers $n_j > m_j$ with the desired property. Q.E.D.

Corollary 4.9 *For all α's, $\#E_\alpha \leq 2$.*

Lemma 4.10 *Notation as in the above lemma. In the case $n.i.(E_\alpha^{(n)}) = 2$ for $n \geq q$, suppose $q \geq 1$ is the minimum integer enjoying that property. Then there are only two possibilities:*

(a) γ_{q-1} is the only escape orbit and $\alpha = 2^{1-q} \pmod{2^{2-q}}$.

(b) γ_n crosses G_{n+1} but not G_{n+2} for all $n \geq q-1$ so that there are two escape orbits and either $\alpha = 2^{2-q}/3 \pmod{2^{2-q}}$ or $\alpha = 2^{3-q}/3 \pmod{2^{2-q}}$.

Proof. First, let us see that γ_{q-1} is the only singular orbit crossing G_q. In fact G_q, by hypothesis, is intersected by only one γ_k $(k \leq q-1)$. (Actually, the case may occur that both γ_p and γ_k cross that aperture, but only if they coincide. Nothing changes in the argument if we take k to be the largest integer of the two.) If $k \leq q-2$ then, by Lemma 4.7, no singular orbit γ_n, with $n \geq k+2$, can "take off". Neither can γ_{k+1}, which would be forced by Lemma 4.6 to pass G_q, against the hypotheses. The net result is that $n.i.(E_\alpha^{(n)}) = 2,\ \forall n \geq k+1$, which contradicts the minimality of q.

Now suppose that γ_{q-1} reaches G_{q+1}. We want to prove that it is also an escape orbit and we are in case *(a)*. In fact, if it stops somewhere after G_{q+1} (say right before $G_k,\ k > q+1$), then γ_q either passes it (and $n.i.(E_\alpha^{(q+1)}) = 3$) or γ_q does not "take off" (and $n.i.(E_\alpha^{(k)}) = 1$). Let us see for which directions this case occurs: from (4.15), $\alpha \in \{2^{-q}, 2^{1-q}\} \pmod{2^{2-q}}$; if $\alpha = 2^{-q} \pmod{2^{2-q}}$, then $n.i.(E_\alpha^{(q)}) = 1$, so that it must be $\alpha = 2^{1-q} \pmod{2^{2-q}}$. Considering $\{E_\alpha^{(n)}\}$, it is easy to see that it consists of two nested sequences of right-open intervals. One of the sequences collapses into the empty set, since all of the intervals share their right endpoint.

So the remaining case is: γ_{q-1} reaches G_q but not G_{q+1}. We would like to prove that this also occurs for all $n > q-1$, i.e., we are in case *(b)*. With the same arguments as above, one checks that either γ_q reaches G_{q+1}, but not G_{q+2}, or it escapes to ∞. The latter cannot be the case, since we already know the only direction for which this can happen, namely $\alpha = 2^{-q} \pmod{2^{1-q}}$: this is the direction for which γ_{q-1} and γ_q coincide, contrary to our present assumption. Reasoning inductively, we obtain the assertion.

Here, as before, we see that $\{E_\alpha^{(n)}\}$ consists of two nested sequences of right-open intervals. But now the two intervals, for a given n, share alternatively (in n) the right and the left endpoint, so that each sequence shrinks to one point. It remains to find the directions corresponding to this case. In the sequel, without loss of generality, we assume $q = 1$.

Let A_n be the set of directions along which γ_n crosses G_{n+1}. According to (4.14), $A_n = \bigcup_{k\in\mathbb{N}}([2^{-n-1}, 3\,2^{-n-1}[+k\,2^{1-n})$. We claim that $A = \bigcap_{n\geq 0} A_n$ is the set of α's we are looking for. In fact, if $\alpha \in A$ then γ_n crosses G_{n+1} for all $n \geq 0$, by definition of A. Moreover, γ_n does not cross G_{n+2} because, if it did, then, by Lemma 4.7, γ_{n+2} would not cross G_{n+3}, which is a contradiction. We note that every A_n has a periodic structure whose fundamental pattern has length 2^{1-n}. The least common multiple of these numbers is 2. Thus, as in the proof of Lemma 4.6, we only need to look at $\widehat{A} := A \cap [0,2]$. This set consists of two points: α_1 and α_2. In fact, let $\widehat{A}_p := [0,2] \cap (\bigcap_{n=0}^{p} A_n)$. Then, referring to Fig. 4.2, it is clear that $\{\widehat{A}_p\}$ is made of two sequences of nested intervals, both having a limit α_i. Furthermore, by the symmetry of the A_n's, $\alpha_2 = 2-\alpha_1$. As indicated by Fig. 4.2, one way to find α_1, and therefore α_2, is to compute the limit of the oscillating sequence $2^{-1}, 2^{-1}+2^{-2}, 2^{-1}+2^{-2}-2^{-3}, 2^{-1}+2^{-2}-2^{-3}+2^{-4}, \ldots$ In other words, $\alpha_1 = \sum_{j=0}^{\infty} 2^{-1-2j} = 2/3$ so that $\alpha_2 = 4/3$. Hence, for $q = 1$, the directions that generate the behavior described in *(b)* are $\alpha = 2/3 \pmod 2$ and $\alpha = 4/3 \pmod 2$. Q.E.D.

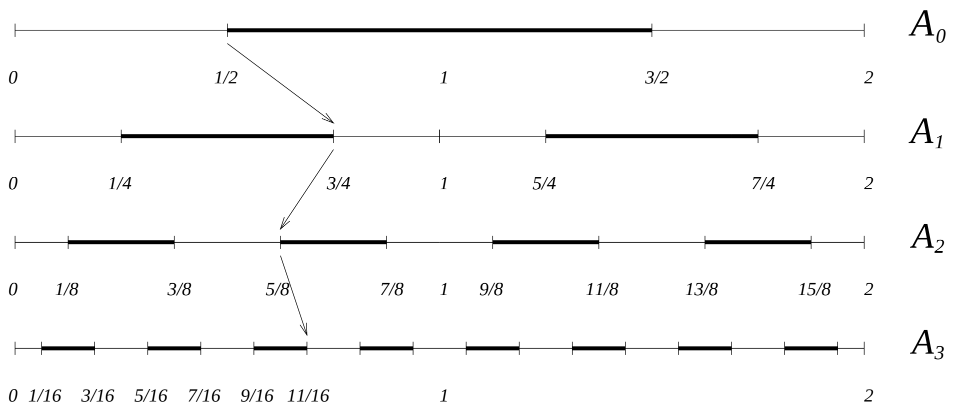

Figure 4.2: The structure of the sets A_n, as in the proof of Lemma 4.10: $A = \bigcap_{n \geq 0} A_n$ consists of two points, both limit of a sequence of nested intervals.

Lemma 4.10 states in particular that, of the two possible cases encountered in Lemma 4.8, the latter occurs for countably many directions. In other words, we have derived for the exponential billiard an even stronger assertion than Theorem 3.1.

We give now our last result concerning $\#E_\alpha$.

Proposition 4.11 *There are no α's without escape orbits.*

Proof. From the previous lemmas, there may be zero escape orbits only for those α's such that there is a sequence $\{n_j\}$ with $n.i.(E_\alpha^{(n_j)}) = 1$. Moreover, in order to have no escape orbits, the intervals $E_\alpha^{(n_j)}$ must eventually share their right extremes. This implies that γ_{n_j} connects the vertex V_{n_j} to the "upper copy" of $V_{n_{j+1}}$, as illustrated in Fig. 3.1. Note that if $n_{j+1} - n_j = 1$ for some $j \geq 0$, then γ_j is an escape orbit (essentially the case *(a)* of Lemma 4.10). We can assume that $n_0 = 0$, otherwise can always rescale the billiard. Thus γ_0 connects the vertices V_0 to V_{n_1}. By looking at (4.12)-(4.13), this happens if, and only if,

$$\alpha = \frac{1}{2} + \frac{1}{n_1 2^{n_1}} + 2k_1 \qquad \text{or} \qquad \alpha = 1 + \frac{1}{n_1 2^{n_1}} + 2k_1, \tag{4.23}$$

for some integer k_1. Now, let us consider γ_1. If we rescale vertically the billiard by a factor 2^{n_1}, we get the same setting we had for γ_0. Since γ_1 connects V_{n_1} to V_{n_2}, we must have, for some k_2,

$$2^{n_1}\alpha = \frac{1}{2} + \frac{1}{(n_2 - n_1)2^{n_2}} + 2k_2 \qquad \text{or} \qquad 2^{n_1}\alpha = 1 + \frac{1}{(n_2 - n_1)2^{n_2}} + 2k_2. \tag{4.24}$$

Since $n_2 - n_1 > 1$ and $n_1 > 1$, a comparison between (4.23) and (4.24) shows that $1/n_1$ must equal $1/2 + 1/((n_1 - n_2)2^{n_2})$ or $1 + 1/((n_1 - n_2)2^{n_2})$. It is not hard to see that this cannot be the case. Therefore there are no α's such that γ_0 intersects V_{n_1} and V_{n_2} at the same time. This proves the statement. Q.E.D.

Lemma 4.8, Lemma 4.10 and Proposition 4.11, can be summarized into the following theorem.

Theorem 4.12 *With reference to the exponential step billiard P: If there exists a non-negative integer m such that $2^m\alpha = 4/3 \pmod 2$, then there are two escape orbits; otherwise there is only one escape orbit.*

We observe that, if there are two escape orbits, they are distinct even in P, by using the same considerations as in the proof of Lemma 4.3.

Remark. The techniques presented in this section to study the escape orbits of the billiard $p_n = 2^{-n}$ can also be applied to the more general case $p_n = 2^{-kn}$, where k is any positive integer. Indeed, for $k > 1$, the analysis of the directions (4.14)-(4.15) along which the semi-orbit γ_p reaches aperture G_m, gives the same qualitative answers as for

$k = 1$. Also, Lemma 4.8, Lemma 4.10 and Proposition 4.11 still hold true, yielding Theorem 4.12.

We finish this chapter by proving the analogue of Lemma 3.4 for the model at hand.

Lemma 4.13 *In the exponential billiard, for a.e. α, η_α does not intersect any vertex.*

PROOF. Suppose there is an α for which the assertion does not hold. If a vertex is contained in the forward semi-orbit, that is, after the material point has crossed L for the last time, then this can only be a singular vertex; hence η_α must contain γ_p, for some $p > 0$. This implies that $n.i.(E_\alpha^{(n)}) = 2$, for n large (see proofs of Lemma 4.5 and 4.2). This case occurs for a null-measure set of directions.

Then assume that some vertex V is contained in the past semi-orbit: V can be $(0, 0)$ or a point of the form $(p, \pm 2^{-q})$ (in particular, Fig. 2.1 shows that $q = |p|$ or $|p| + 1$). Let us call $(0, Y_0)$ the last (in time) intersection point between the orbit and L. Let j be the number of rectangular boxes visited by the orbit after leaving V and before reaching $(0, Y_0)$. Then, in complete analogy with (4.10),

$$Y_0 = m_0 2^{-q} + \sum_{i=1}^{j} (\alpha - m_i 2^{-q_i}); \tag{4.25}$$

with $m_i, q_i \geq 0$ some integers ($m_0 = 0$ or ± 1 depending on V being the origin or not).

We turn now to the rescaled coordinates: $y_0 = Y_0/2$. Rearranging the previous equality yields, for some integers m, k,

$$y_0 = m 2^{-k} + \frac{j}{2}\alpha. \tag{4.26}$$

Since the orbit is supposed to escape after leaving $(0, Y_0)$, we can apply (4.4), (4.5) with y_0 as in (4.26). If $n \geq k$ the first term in (4.26) gets canceled. Therefore one must have

$$y_{n+1} = (n + j + 1) 2^n \alpha \pmod 2 \in \left[-\frac{1}{2}, \frac{1}{2} \right[\quad \forall n \geq k. \tag{4.27}$$

Define the increasing sequence $\{n_i\}_{i \geq p}$ by $n_i + j + 1 = 2^i$ with $n_p \geq k$. Condition (4.27) implies in particular that

$$y_{n_i+1} = 2^{i+n_i} \alpha \pmod 2 \in \left[-\frac{1}{2}, \frac{1}{2} \right[\quad \forall i \geq p. \tag{4.28}$$

One verifies that—except for countably many α's whose binary expansion in not uniquely defined—(4.28) is equivalent to saying that the $-(i + n_i + 1)$-th digit of the binary expansion of α is a zero for every $i \geq p$ (cf. Appendix of [DDL1]). The Lebesgue measure makes these events independent and equally likely with probability 1/2. Hence (4.28) can only occur for a null-measure set of directions. This proves that for almost no α's, η_α can start from vertex V and pass through j boxes before taking off to infinity. Since events like this are countably many, we have shown that an escape orbit can almost never contain a vertex in its backward part. Q.E.D.

Chapter 5

Ergodic properties

One would like to know about the ergodic properties of the infinite step billiards. The statements of Chapter 3, although fairly interesting from a topological point of view, do not address at all the measure-theoretic features of these dynamical systems. (For example, one should not be misled by Corollary 3.8,*(iii)*, which is a much weaker statement than ergodicity.)

Something can be done in this direction, starting from the available results for finite polygonal tables. Since the *typical* truncated billiard is ergodic (Proposition 1.1), one can approximate an infinite step polygon P by means of a suitable $P^{(n)}$ (recall the definition from Chapter 2). If the approximation is so good that the dynamics on the two billiards are very similar for most initial points and for long times, then one can use the ergodicity of the truncated table to obtain the same result for P.

This is certainly not a new idea, being essentially the argument behind the main statements of Proposition 1.3 [KMS]. Recently Troubetzkoy [Tr] revived it to prove that, in a certain space of generic infinite polygons, the typical element is ergodic for almost every direction θ. In Section 5.2 we do the same for our more modest class of tables. (The curious reader will find in [DDL2] more comments about the analogies and differences between our Theorem 5.2 and [Tr, Thm. 5.1].)

Before moving on to that, we present in Section 5.1 a somewhat weaker theorem which shows that there are infinite step billiards ergodic in one given direction. The purpose of this is to give a more readable and constructive proof than that of Theorem 5.2.

Completing the chapter is a section regarding the entropy of the billiard w.r.t. certain ergodic measures.

5.1 Ergodicity in one direction

Theorem 5.1 *Fix $\alpha \notin \mathbb{Q}$. For every positive, decreasing, vanishing sequence $\{\bar{p}_n\} \subset \mathbb{Q}$, and every integer k, there exists a decreasing sequence $\{p_n\} \subset \mathbb{Q}$, with*

$$\begin{aligned} p_n = \bar{p}_n, &\qquad \text{for } 0 \le n \le k; \\ 0 < p_n \le \bar{p}_n, &\qquad \text{for } n > k; \\ \sum_n p_n < \infty. & \end{aligned}$$

such that the billiard flow $\phi_{\alpha,t}$ on S_α^P, for $P \simeq \{p_n\}$, is ergodic (hence almost all orbits are dense).

PROOF OF THEOREM 5.1. We will construct S_α in such a way that almost every point in L has a typical trajectory, in the sense that the time average w.r.t. ϕ_t of a function in a dense subspace of $L^1(S_\alpha)$ equals its spatial average. Since L is a Poincaré section, the same property will hold for a.e. point in S_α, thus proving that ϕ_t is ergodic.

For the sake of notation, we will drop the subscript α in the sequel. It may be useful to remark that the Lebesgue measure, that we use here, is not normalized.

Take a positive sequence $\varepsilon_n \searrow 0$. Then set $p_i := \bar{p}_i \; \forall i \le k$. We are going to build the rest of our billiard by induction: suppose we have fixed p_i for $1 \le i \le n$, $(n \ge k)$ and we have to determine a suitable p_{n+1}. Consider $S^{(n)}$, generated by the p_i's found so far. The flow $\phi_t^{(n)}$ on it is ergodic by Proposition 1.1,*(i)* and *(iii)*. For $f \in L^1(S^{(n)})$ and $z \in L$ define

$$\left(\Xi_{(n)}^T f\right)(z) := \frac{1}{T}\int_0^T f \circ \phi_t^{(n)}(z)\, dt - \frac{1}{|S^{(n)}|}\int_{S^{(n)}} f \, dX dY. \tag{5.1}$$

Let $\{f_j^{(n)}\}_{j\in\mathbb{N}}$ be a separable basis of $L^1(S^{(n)})$. For the rest of the proof $S^{(n)}$ will be naturally identified with an open subset of $S^{(m)}$, $m > n$ (see Fig. 2.2). As a consequence, a function defined on the former set will be implicitly extended to the latter by setting it null on the difference set. With this in mind, let

$$A_T^{(n)} := \left\{ z \in L \,|\, \forall\, 1 \le i, j \le n, \left| \Xi_{(n)}^T f_i^{(j)}(z) \right| \le \varepsilon_n \right\}. \tag{5.2}$$

By ergodicity, since only a finite number of functions are involved in the above set, we have $|A_T^{(n)}| \to |L| = 2$ as $T \to \infty$. Take T_n such that $|A_{T_n}^{(n)}| \ge 2 - \varepsilon_n/2$. We are now in position to determine p_{n+1}. Choose some

$$p_{n+1} \in \mathbb{Q}; \quad 0 < p_{n+1} \le \min \left\{ \bar{p}_{n+1}, \frac{\varepsilon_n}{2T_n} \right\} \tag{5.3}$$

(this can be done in such a way that eventually $\sum_n p_n < \infty$). Now imagine to open a hole of width $2p_{n+1}$ in the middle of $\{n+1\} \times [-p_n, p_n[$ (same as $\{-n-1\} \times [-p_n, p_n[$ since they are identified at the moment). The motion on $S^{(n)}$ is not affected very much by this change, during the time T_n. If we denote by ϕ_t the flow on the infinite billiard table (when we are done constructing it), we can already assert that, taken a point $z \in L$, $\phi_t^{(n)}(z) = \phi_t(z)\, \forall t \in [0, T_n]$ unless the particle departing form z hits the hole in a time less than T_n. We can estimate the measure of these "unlucky" initial points: they constitute the set

$$B_n := L \cap \left(\bigcup_{t \in [-T_n, 0]} \phi_t^{(n)}(\{n+1\} \times [-p_{n+1}, p_{n+1}[\,) \right). \tag{5.4}$$

The backward beam (up to time $-T_n$) originating from the hole cannot hit L more than $T_n/2$ times, since between each two successive crossings of L, the beam has to cover a distance which is at least 2 (see Fig. 2.2), but the velocity of the particles was conventionally fixed to 1. Every intersection of the beam with L is a set of measure $2p_{n+1}$, so, from (5.3), $|B_n| \le \varepsilon_n/2$.

Set $C_n := A_{T_n}^{(n)} \backslash B_n$, thus $|C_n| \ge 2 - \varepsilon_n$. So C_n is the set of points which keep enjoying the properties as in (5.2), even after the cut has been made in $S^{(n)}$. Suppose one recursively defines $p_n\, \forall n$, thus determining an infinite manifold S. Let $C := \bigcap_{n\in\mathbb{N}} \bigcup_{m \ge n} C_m$. Then $|C| = \lim_{n\to\infty} \cup_{m\ge n} C_m = 2 = |L|$. This is the "good" set since, fixed $z \in C$, there exist a subsequence $\{n_k\}$ such that $z \in \bigcap_k C_{n_k}$. This means that, taken two integers i, j, $\forall n_k \ge \max\{i, j\}$,

$$\left| \frac{1}{T_{n_k}} \int_0^{T_{n_k}} f_i^{(j)} \circ \phi_t(z)\, dt - \frac{1}{|S^{(n_k)}|} \int_S f_i^{(j)}\, dX dY \right| \le \varepsilon_{n_k}. \tag{5.5}$$

Comparing this with (5.1) we notice two differences. First, the flow that appears here is ϕ_t because of the remark after (5.3). Second, the manifold integral is taken over all of S: this is so because of the initial convention to extend with zero all functions defined on subsets of S.

Define Ξ^T in analogy with (5.1). Since $|S^{(n)}| \nearrow |S|$, (5.5) shows that $(\Xi^{T_{n_k}} f_i^{(j)})(z) \to 0$, as $k \to \infty$, with T_n in general going to ∞ (this is not guaranteed by the definition of T_n, but one can easily arrange to make this happen). We now invoke Birkhoff's Theorem, which states that, for the function $f_i^{(j)} \in L^1(R)$, the time average is well-defined a.e. (in R, hence in L). Therefore, for every $f \in \mathrm{span}\{f_i^{(j)}\}_{i,j\in\mathbb{N}}$, there exists a set $C_f \subseteq L$, $|C_f| = 2$ such that

$$\lim_{T\to+\infty} (\Xi^T f)(z) = 0. \tag{5.6}$$

This proves the claim we made at the beginning, since $\mathrm{span}\{f_i^{(j)}\}$ is by construction dense in $L^1(S)$.

As concerns the assertion about the density, that immediately follows from standard arguments as in [W, Thm. 5.15] (which can be checked to hold under our hypotheses, as well). Q.E.D.

Remark. The fact that the above result provides ergodic billiards with rational heights only is merely technical. Had we decided not to use almost integrable billiards, we could not have exploited Proposition 1.1,*(iii)*, but the result would have followed anyway, in a slightly more complicated fashion.

5.2 Generic ergodicity

Let $\mathcal{S}$ be the space of all step polygons with unit area. (This means that, in this section, we drop the unsubstantial requirement $p_0 = 1$.) Since each step polygon P is uniquely determined by the sequence $\{p_n\}_{n\in\mathbb{N}}$, we apply to $\mathcal{S}$ the metric of the space ℓ^1: given $P \simeq \{p_n\}$, $Q \simeq \{q_j\}$, let

$$d(P,Q) := \sum_{n=0}^{\infty} |p_n - q_n|. \tag{5.7}$$

The metric space $(\mathcal{S}, d)$ is complete and separable. Note that $d(P,Q) = A(P \triangle Q)$. Also, denote by $\mathcal{S}_0$ the collection of all finite $P \in \mathcal{S}$ with rational heights.

The following statements are the highlight of this section:

Theorem 5.2 *For a typical step polygon P, the flow on S_α^P is ergodic for a.e. $\alpha \in \mathbb{R}^+$.*

Corollary 5.3 *There are infinite step billiards P ergodic on S^P_α for a.e. α.*

In view of the proofs, we turn to the notation $\theta = \arctan\alpha$ to denote directions. More specifically, we use the normalized Lebesgue measure on $\theta \in]0, \pi/2[$. Although this is equivalent (in the sense of reciprocal absolute continuity) to the Lebesgue measure on $\alpha \in \mathbb{R}^+$, the former will simplify the arguments quite a bit.

We need some definitions and a technical lemma. In $\mathbb{R}^2$, let ρ be the Euclidean metric and A the area. Also, denote by $\mathcal{V} := \{V_k\}$ the set of non-removable singularities of S^P_θ, as defined at the beginning of Chapter 2.

Definition 5.4 *Given $P, Q \in \mathcal{S}$, $\theta \in]0, \pi/2[$, $x \in S^P_\theta \cap S^Q_\theta$ and $\varepsilon > 0$, let*

$$I(P, Q, x, \varepsilon) := \{t \in [0, 1/\varepsilon] \mid \rho(\phi^P_{\theta,t}(x), \phi^Q_{\theta,t}(x)) > \varepsilon\}, \tag{5.8}$$

$$G(P, Q, \theta, \varepsilon) := \{x \in S^P_\theta \cap S^Q_\theta \mid |I(P, Q, x, \varepsilon)| < \varepsilon\}, \tag{5.9}$$

$$E(P, Q, \varepsilon) := \left\{\theta \in]0, \pi/2[\mid A(G(P, Q, \theta, \varepsilon)) \geq A(S^{P,Q}_\theta) - \varepsilon\right\}. \tag{5.10}$$

Lemma 5.5 *For any $P \in \mathcal{S}$ and $\varepsilon > 0$, we have*

$$\lim_{Q \to P} |E(P, Q, \varepsilon)| = 1.$$

Proof of Lemma 5.5. It is enough to prove the statement for any sequence converging to P. Let $\{P_n\}$ be such a sequence. Fix $\theta \in]0, \pi/2[$ and $\varepsilon > 0$. Only countably many orbits of S^P_θ contain a singular vertex V. Let $x \in S^P$ be a point that does not belong to any of these orbits. In a finite interval of time its trajectory can get close only to a finite number of singular vertices. Therefore δ_ε, the distance between the $\cup_{0 \leq t \leq 1/\varepsilon} \phi^P_{\theta,t}(x)$ and $\mathcal{V}$, is positive. Let N_ε be the number of collisions of $\cup_{0 \leq t \leq 1/\varepsilon} \phi^P_{\theta,t}(x)$ with the horizontal sides of S^P_θ.

The two surfaces S^P_θ and $S^{P_n}_\theta$ overlap (as subsets of $\mathbb{R}^2$). If $x \in S^{P_n}$ as well, then the set $\cup_{0 \leq t \leq 1/\varepsilon} \phi^{P_n}_{\theta,t}(x)$ represents a finite piece of the trajectory of x in the surface $S^{P_n}_\theta$. We want to estimate $\rho(\phi^P_t(x), \phi^{P_n}_t(x))$ for $0 \leq t \leq 1/\varepsilon$. Since there is a natural correspondence between the sides of the two surfaces, when we say that the two trajectories hit the same sequence of sides, we mean corresponding sides. Notice that $\rho(\phi^P_t(x), \phi^{P_n}_t(x))$ stays constant when neither orbit crosses any sides. Also, as long as $\phi_t^P(x)$ and $\phi_t^{P_n}(x)$ hit the same sequence of sides, every time that there is collision at a horizontal side, the distance $\rho(\phi_t^P(x), \phi_t^{P_n}(x))$ increases by a term $h \leq 2d(P, P_n)$. Hence the maximum distance between the two trajectories in the interval $t \in [0, 1/\varepsilon]$ is $\leq 2d(P, P_n)N_\varepsilon$. Therefore, by definition of δ_ε, if $d(P, P_n) < \delta_\varepsilon/(2N_\varepsilon)$, the trajectories encounter the same sequence of sides for $0 \leq t \leq 1/\varepsilon$.

Since $d(P, P_n) \to 0$ as $n \to \infty$, we can find an $m(x, \varepsilon) > 0$ such that the previous inequality is satisfied for all $n > m(x, \varepsilon)$. It is clear now that, as n grows larger, the distance between the two trajectories decreases. We conclude that, for points $x \in S_\theta^P$ with a non-singular positive semi-trajectory,

$$\lim_{n\to\infty} \max_{0\leq t\leq 1/\varepsilon} \rho(\phi_t^P(x), \phi_t^{P_n}(x)) = 0. \tag{5.11}$$

As a consequence, we have that a.e. $x \in S_\theta^P$ belongs to $G(P, P_n, \theta, \varepsilon)$ if n, which depends on x, is sufficiently large. Therefore $A(G(P, P_n, \theta, \varepsilon)) \to A(S_\theta^P)$ as $n \to \infty$. Being this true for all $\theta \in]0, \pi/2[$, we finally obtain $|E(P, P_n, \varepsilon)| \to 1$ as $n \to \infty$, for any $\varepsilon > 0$. Q.E.D.

We are now in position to attack the main proof of this section:

Proof of Theorem 5.2. Choose $\varepsilon_n > 0$ such that $\lim_{n\to\infty} \varepsilon_n = 0$. Let $\{f_i\}_{i\in\mathbb{N}}$ be a countable collection of continuous functions with compact support and let it be dense in $L^1(\mathbb{R}^2)$. For each step polygon P, f_i^P denotes the restriction of f_i to S_θ^P, corrected to obey the identifications on S_θ^P. These corrections occur on a set of zero Lebesgue measure in $\mathbb{R}^2$, therefore $\{f_i^P\}_{i\in\mathbb{N}}$ is dense in $L^1(S_\theta^P)$ for each $\theta \in]0, \pi/2[$.

Given $P \in \mathcal{S}$, $x \in S_\theta^P$ and $i \in \mathbb{N}$, let us introduce

$$B_T^P(\theta, i, x) := \left| \frac{1}{T} \int_0^T f_i^P \circ \phi_t^P(x) dt - \frac{1}{A(S_\theta^P)} \int_{S_\theta^P} f_i^P dA \right| \tag{5.12}$$

and

$$C_T^P(\theta, n) := \left\{ x \in S_\theta^P \mid B_T^P(\theta, i, x) \leq \varepsilon_n,\ i = 1, \ldots, n \right\}. \tag{5.13}$$

If $P \in \mathcal{S}_0$, the billiard flow $\phi_{\theta,t}^P$ is uniquely ergodic for all irrational θ (i.e., $\alpha = \tan\theta \notin \mathbb{Q}$). Thus, for every $n > 0$ and such θ, we have that $\lim_{T\to+\infty} A(C_T^P(\theta, n)) = A(S_\theta^P)$. Let

$$D_T^P(n) := \left\{ \theta \in]0, \pi/2[\mid A(C_T^P(\theta, n)) \geq A(S_\theta^P) - \varepsilon_n \right\}. \tag{5.14}$$

Then we can choose a $T_n(P) \geq n$ such that $|D_T^P(n)| > 1-\varepsilon_n$ for any $T \geq T_n(P)$. Since $f_1, f_2, \ldots, f_n$ are uniformly continuous on $\mathbb{R}^2$, there is an $r_n > 0$ for which $|f_i(x) - f_i(y)| \leq \varepsilon_n$ whenever $\rho(x,y) \leq r_n$ and $i = 1, \ldots, n$. Let

$$\delta_n(P) := \min\left\{\frac{1}{T_n(P)}, \varepsilon_n, \frac{\varepsilon_n}{\max_{1\leq i\leq n}\|f_i\|_\infty}, r_n\right\} \tag{5.15}$$

and $\tau_n := 1/\delta_n$. By Lemma 5.5, there exists $0 < \sigma_n(P) \leq \delta_n(P)$ such that, if $Q \in U_{\sigma_n}(P) := \{R \in \mathcal{S} \mid d(P,R) \leq \sigma_n\}$, then $|E(P,Q,\delta_n)| > 1-\delta_n \geq 1-\varepsilon_n$.

Let $E_n := E(P,Q,\delta_n)$ and $I_n(x)$, $G_n(\theta)$ be the sets (5.8), (5.9) used in the definition of $E(P,Q,\delta_n)$. These sets depend on P. For $\theta \in E_n$ and $x \in G_n(\theta)$, $A(G_n(\theta)) \geq A(S_\theta^{P,Q}) - \delta_n \geq A(S_\theta^{P,Q}) - \varepsilon_n$, $|I_n(x)| < \delta_n$ and $\rho(\phi_t^P(x), \phi_t^Q(x)) \leq \delta_n \leq r_n$ for $t \in [0,\tau_n] \setminus I_n(x)$. Notice that $T_n \leq \tau_n$. For $i = 1, \ldots, n$, we have:

$$\begin{aligned}
&\left|\frac{1}{A(S^P)}\int_{S^P} f_i^P dA - \frac{1}{A(S^Q)}\int_{S^Q} f_i^Q dA\right| \\
&\quad = \frac{1}{A(S^P)}\left|\int_{S^P} f_i^P dA - \int_{S^Q} f_i^Q dA\right| \\
&\quad \leq \frac{1}{4}\int_{S^P\triangle S^Q} |f_i|\, dA \leq \frac{1}{4}\|f_i\|_\infty A(S^P\triangle S^Q) \\
&\quad \leq \sigma_n\|f_i\|_\infty \leq \delta_n\|f_i\|_\infty \leq \varepsilon_n.
\end{aligned} \tag{5.16}$$

Let $\theta \in D_{\tau_n}^P(n) \cap E_n$ and $x \in C_{\tau_n}^P(\theta,n) \cap G_n(\theta)$. Then

$$\begin{aligned}
&\left|\frac{1}{\tau_n}\int_0^{\tau_n} f_i^Q \circ \phi_t^Q(x)dt - \frac{1}{A(S^Q)}\int_{S^Q} f_i^Q dA\right| \\
&\quad \leq \frac{1}{\tau_n}\int_0^{\tau_n}\left|f_i^Q \circ \phi_t^Q(x) - f_i^P \circ \phi_t^P(x)\right| dt \\
&\quad + \left|\frac{1}{\tau_n}\int_0^{\tau_n} f_i^P \circ \phi_t^P(x)dt - \frac{1}{A(S^P)}\int_{S^P} f_i^P dA\right| \\
&\quad + \left|\frac{1}{A(S^P)}\int_{S^P} f_i^P dA - \frac{1}{A(S^Q)}\int_{S^Q} f_i^Q dA\right| \\
&\quad =: I + II + III.
\end{aligned} \tag{5.17}$$

We have $I \leq 2\,\delta_n^2\|f_i\|_\infty + \varepsilon_n \leq 3\varepsilon_n$—say—for n large enough. Moreover, $x \in C_{\tau_n}^P(\theta,n)$ implies $II \leq \varepsilon_n$. Finally, $III \leq \varepsilon_n$ by (5.16). We conclude that $I + II + III \leq 5\varepsilon_n$ for $\theta \in D_{\tau_n}^P(n) \cap E_n$, $x \in C_{\tau_n}^P(\theta,n) \cap G_n(\theta)$ and $i = 1, \ldots, n$. By definition of $D_{\tau_n}^P(n)$ and E_n, we have

$$|D_{\tau_n}^P(n) \cap E_n| > 1 - 2\varepsilon_n. \tag{5.18}$$

From $A(C^P_{\tau_n}(\theta, n)) \geq A(S^P_\theta) - \varepsilon_n$ and $A(G_n(\theta)) \geq A(S^P_\theta) - \varepsilon_n$ it follows that

$$A(C^P_{\tau_n}(\theta, n) \cap G_n(\theta)) \geq A(S^P_\theta) - 2\varepsilon_n = A(S^Q_\theta) - 2\varepsilon_n. \tag{5.19}$$

Let $\{P_j\}_{j \in \mathbb{N}}$ be an enumeration of $\mathcal{S}_0$ and

$$H := \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} U_{\sigma_n(P_j)}(P_j). \tag{5.20}$$

It is easy to see that H is a dense G_δ of $\mathcal{S}$. If $Q \in H$, then for every $n > 0$ there is a j_n for which $Q \in U_{\sigma_n(P_{j_n})}(P_{j_n})$. Define $D := \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} D^{P_{j_n}}_{\tau_n}(n) \cap E_n$. By (5.18), $|D| = 1$. This means that, for each $\theta \in D$, there is a subsequence $\{n_k\}$ such that $\theta \in D^{P_{j_{n_k}}}_{\tau_{n_k}}(n_k)$ for all k. In order to avoid heavy notation, let us denote such a sequence by $\{n\}$. Now call $C(\theta) := \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} C^{P_{j_n}}_{\tau_n}(\theta, n) \cap G_n(\theta)$. From (5.19), $A(C(\theta)) = A(S^Q_\theta)$.

So, for each $\theta \in D$ and $x \in C(\theta)$ (i.e., a.e. θ and a.e. $x \in S^Q_\theta$) there exists a subsequence $\{n_k\}$ such that $\lim_{k\to\infty} \tau_{n_k} = +\infty$ and

$$\lim_{k\to\infty} \frac{1}{\tau_{n_k}} \int_0^{\tau_{n_k}} f^Q_i \circ \phi^Q_t(x) dt = \frac{1}{A(S^Q)} \int_{S^Q} f^Q_i dA, \tag{5.21}$$

for all $i \in \mathbb{N}$. Since $\{f^Q_i\}_{i \in N}$ is dense in $L^1(S^Q_\theta)$, using a standard approximation argument, and Birkhoff's Theorem, we conclude that $\phi^Q_{\theta,t}$ is ergodic for a.e. $\theta \in]0, \pi/2[$. Q.E.D.

Proof of Corollary 5.3. Let $F^N \subset \mathcal{S}$ be the collection of finite step polygons with N sides. Each F^N is a nowhere dense set, so the family of all truncated step polygons $F := \bigcup_{N \geq 4} F^N$ is a subset of first Baire's category (sometimes this is referred to as a *meager* set) in $\mathcal{S}$. By Baire's Theorem, H is a set of second category, therefore it intersects the complement of F. In other words, H contains truly infinite step polygons. Q.E.D.

5.3 Metric entropy

Let us go back to our usual notation α for the directions. We prove the following property of non-periodic orbits, which turns out to be very useful in certain entropy estimates, later in this section.

Proposition 5.6 *The closure of the forward and backward semi-orbit of every non-periodic point intersects* $\mathcal{V} \cup \{\infty\}$.

Remark. This result can be considered an adaptation of Proposition 1.2,*(ii)*. Notice that in our case the billiard is always rational (or weakly rational according to [Tr, Def. 4.3]) and $\mathcal{V}$ is only a subset of all vertices.

Proof of Proposition 5.6. If a (past or future) semi-orbit γ of a non-periodic point is bounded, then it is entirely contained in a finite step billiard $P^{(n)}$, for n large enough. According to [GKT], the closure of γ must contain a singular vertex of $P^{(n)}$, which clearly is in $\mathcal{V}$. If instead γ is unbounded, then the set of its limiting points contains $\{\infty\}$. Q.E.D.

Take $n \geq 1$. In Fig. 2.1, let $L_n := \{n\} \times [p_n, p_{n-1}[$ and $L_{-n} := \{n\} \times [-p_{n-1}, -p_n[$ be the two copies of the n-th vertical side. We identify them with $d_n := [p_n, p_{n-1}[$ and $d_{-n} := [-p_{n-1}, -p_n[$. The family of these intervals partitions $I := [-1, 0[\cup]0, 1[$. Thus it makes sense to define f_α as the i.e.t. induced by $\phi_{\alpha,t}$ on I.

For any f_α-invariant Borel probability measure ν, we set

$$H_\nu(P) := -\sum_{n \neq 0} \nu(d_n) \log \nu(d_n). \tag{5.22}$$

Call $\widetilde{\mathcal{X}}_\alpha := L \setminus \bigcap_{n=0}^{\infty} \mathcal{P}_\alpha^{-n} E_\alpha$ the set of the points in L whose forward orbits keep returning there; and denote by $\mathcal{X}_\alpha$ the corresponding set in I. More precisely,

$$\mathcal{X}_\alpha := g_\alpha\left(\widetilde{\mathcal{X}}_\alpha\right), \tag{5.23}$$

where $g_\alpha : \widetilde{\mathcal{X}}_\alpha \longrightarrow I$ is given by $g_\alpha(x) = \phi_{\alpha,t_1}(x)$ and t_1 is the first collision time at a vertical wall. For any $x \in \widetilde{\mathcal{X}}_\alpha$, let

$$\widetilde{\pi}_\alpha(x) := \{\ldots, L_{\omega_{-1}}, L, L_{\omega_0}, L, L_{\omega_1}, L, \ldots, L, L_{\omega_k}, \ldots\}, \tag{5.24}$$

be the sequence of vertical sides that the orbit of x crosses (adopting the convention that L_{ω_0} is the first vertical side encountered in the past). Then we define the *coding* $\pi_\alpha : \widetilde{\mathcal{X}}_\alpha \longrightarrow (\mathbb{Z}^*)^{\mathbb{Z}}$ by

$$\pi_\alpha(x) := \{\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots\} \tag{5.25}$$

We equip $C_\alpha := \pi_\alpha(\mathcal{X}_\alpha)$ with the product topology and we denote by σ the left shift. The following diagram commutes:

$$
\begin{array}{rcl}
f_\alpha : \mathcal{X}_\alpha & \longrightarrow & \mathcal{X}_\alpha \\
\uparrow & g_\alpha & \uparrow \\
\mathcal{P}_\alpha : \widetilde{\mathcal{X}}_\alpha & \longrightarrow & \widetilde{\mathcal{X}}_\alpha \\
\downarrow & \pi_\alpha & \downarrow \\
\sigma : C_\alpha & \longrightarrow & C_\alpha
\end{array}
\tag{5.26}
$$

Proposition 5.7 *Let ν be any f_α-invariant, Borel, ergodic probability measure on $\mathcal{X}_\alpha$. If $H_\nu(P) < \infty$, then $h_\nu = 0$.*

Remark. For example, if p_n decays exponentially or faster, as for the tables in Chapter 3 and 4, then $H_\ell(P)$ is finite for the Lebesgue measure ℓ.

Proof of Proposition 5.7. First of all, notice that the assertion is obvious if ν is atomic. We claim that the partition $\{A_1, A_2, \ldots\}$ given by

$$A_j := \{\omega \in C_\alpha \,|\, \omega_1 \;=\; j\} \tag{5.27}$$

is a one-side generator. This implies by standard results [R, 11.3] that $h_\lambda = 0$ for any σ-invariant Borel probability measure λ on C_α. Now, g_α is bijective by construction, therefore the pull-back $g_{\alpha *}\nu$ is also non-atomic and ergodic. Hence, there is a $Y \subseteq \widetilde{\mathcal{X}}_\alpha$ such that $\pi_\alpha|_Y$ is injective and $g_{\alpha *}\nu(Y) = 1$. We conclude that $h_\nu(f_\alpha) = h_{\pi^*_\alpha g_{\alpha *}\nu}(\sigma) = 0$.

It remains to prove the initial claim. This is done if we show that the sequence of sides an orbit visits in the future determines the sequence of sides visited in the past. Clearly, if the orbit is periodic, there is nothing to prove. In the non-periodic case we can use Proposition 5.6 to conclude that any parallel strip of orbits must eventually "split" at some vertex V_k and this of course implies that two distinct orbits cannot hit the same vertical sides in the future. Q.E.D.

The topological entropy, too, is probably zero (at least for many of these systems) but, due to the non-compactness of the table, the usual variational principle cannot be applied directly and one must check some additional conditions [PP].

Part II

Large deviations for ideal quantum systems

Chapter 6

Introduction

Statistical mechanics is the bridge between the microscopic world of atoms and the macroscopic world of bulk matter. In particular it provides a prescription for obtaining macroscopic properties of systems in thermal equilibrium from a knowledge of the microscopic Hamiltonian. This prescription becomes mathematically precise and elegant in the limit in which the size of the system becomes very large on the microscopic scale (but not large enough for gravitational interactions between the particles to be relevant). Formally this corresponds to considering neutral or charged particles with effective translation invariant interactions inside a container and taking the infinite-volume or thermodynamic limit (TL). This is the limit in which the volume $|V|$ of the container V grows to infinity along some specified regular sequences of domains, say cubes or balls, while the particle and energy density approach some finite limiting value [R, F, G, T]. This limit provides a precise way for eliminating "finite size" effects.

It is then an important result (a theorem, under suitable assumptions) of statistical mechanics that the bulk properties of a physical system, computed from the thermodynamic potentials via any of the commonly used Gibbs ensembles (microcanonical, canonical, grand canonical, etc.), have well defined "equivalent" TL's [R, F, G]. These free energy densities are furthermore proven to be the same for a suitable class of "boundary conditions" (b.c.), describing the interaction of the system with the walls and the "outside" of its container. When this independence of b.c. is "strong enough", the bulk free energies also yield information about normal fluctuations and large deviations (LD), in particle number and energy, inside regions Λ that are macroscopically large but significantly smaller than V. The restriction to $|\Lambda| \ll |V|$ means that we deal here with *semi-local* rather than global LD.

Part II of this dissertation investigates LD in the particle number, for possibly the simplest quantum systems, free fermions and bosons. This seems interesting since the real world is quantum mechanical, with the classical description being an essentially uncontrolled approximation, albeit a very good one in many circumstances.

The theory of LD for quantum systems is at the present time much less developed than for classical ones [R, G, vFS]. In particular the very interesting works of Lewis and collaborators on Bose systems (e.g., [vLP2, LZP, DLP]) only deal with LD in the full macroscopic box V, i.e., with global fluctuations. There is also a theory of normal and anomalous small fluctuations for semi-local observables, again for some Bose systems: see [GVV, BMV, ABV] and references there.

6.1 Classical systems

We begin by considering a classical system of N particles of mass m in a domain, say a cubical box $V \subset \mathbb{R}^d$, interacting with each other through a sufficiently rapidly decaying pair potential $\phi(r)$, e.g., a Lennard-Jones potential. The Hamiltonian of the system is then given by

$$H(N, V; b) := \frac{1}{2m} \sum_{i=1}^{N} \mathbf{p}_i^2 + \frac{1}{2} \sum_{1 \leq i \neq j \leq N} \phi(r_{ij}) + \sum_{i=1}^{N} u_b(\mathbf{r}_i), \tag{6.1}$$

where $\mathbf{p}_i \in \mathbb{R}^d$, $\mathbf{r}_i \in V$, $r_{ij} := |\mathbf{r}_i - \mathbf{r}_j|$, and $u_b(\mathbf{r}_i)$ represents the interaction of the i-th particle with the world outside of the boundary of V. This boundary interaction (indicated here and in the sequel by b) is in addition to the action of the implicitly assumed "hard wall" which keeps the particles confined to V. The dynamic effect of the latter is to reflect the normal component of the particle's momentum when it hits the wall. However, sometimes it is convenient to replace it with periodic boundary conditions [FL], dropping the boundary term u_b in (6.1).

For a macroscopic system in equilibrium at reciprocal temperature β and chemical potential μ, the grand canonical Gibbs ensemble then gives the probability density for finding exactly N particles inside $V \subset \mathbb{R}^d$ at the phase point $X_N := (\mathbf{r}_1, \mathbf{p}_1, \ldots, \mathbf{r}_N, \mathbf{p}_N) =:$

$(\mathbf{R}_N, \mathbf{P}_N) \in \Gamma_N := V^N \times \mathbb{R}^{dN}$ as

$$\nu(X_N \,|\, \beta, \mu, V, b) := \frac{(N!)^{-1} h^{-Nd} \exp[-\beta(H(N,V;b) - \mu N)]}{\Xi(\beta, \mu \,|\, V, b)}. \tag{6.2}$$

Here Ξ is the grand canonical partition function

$$\begin{aligned}\Xi &:= \sum_{N=0}^{\infty} (N!)^{-1} \lambda_B^{-dN} e^{\beta\mu N} \int_{V^N} d\mathbf{r}_1 \cdots d\mathbf{r}_N \, e^{-\beta/2 \sum \phi(r_{ij}) - \beta \sum u_b(\mathbf{r}_i)} \\ &=: \sum_{N=0}^{\infty} e^{\beta\mu N} \, Q(\beta, N \,|\, V, b),\end{aligned} \tag{6.3}$$

and $Q(\beta, N \,|\, V, b)$ is the canonical partition function. We use h^{dN}, h being Planck's constant, as the unit of volume in the phase space Γ_N, so $\lambda_B := h\sqrt{\beta/(2\pi m)}$ is the de Broglie wave length. The finite-volume, boundary condition dependent, grand canonical pressure is

$$p(\beta, \mu \,|\, V, b) := (\beta|V|)^{-1} \log \Xi(\beta, \mu \,|\, V, b). \tag{6.4}$$

Taking now the TL, $V \nearrow \mathbb{R}^d$, we obtain, for a suitable class of b.c., an intrinsic (b.c. independent) grand canonical pressure $p(\beta, \mu)$. This is related to the Helmholtz free energy density $a(\beta, \rho)$ in the canonical ensemble, obtained when Ξ is replaced by $Q^{-1}(\beta, N \,|\, V, b)$ in (6.4) and the limit is taken in such a way that $N/|V| \to \rho$, a specified particle density. The relation between p and a is given by the usual thermodynamic formula involving the Legendre transform

$$p(\beta, \mu) = \sup_{\rho} [\rho\mu - a(\beta, \rho)] = \pi(\beta, \bar{\rho}), \tag{6.5}$$

where $\pi(\beta, \rho)$ is the TL of the canonical pressure

$$\pi(\beta, \rho) := -\rho^2 \frac{\partial(a/\rho)}{\partial\rho} \tag{6.6}$$

and

$$\bar{\rho}(\beta, \mu) := \frac{\partial p}{\partial \mu}(\beta, \mu) \tag{6.7}$$

is the average density in the grand canonical ensemble.

At a first order phase transition $\mu \mapsto \bar{\rho}(\beta, \mu)$ is discontinuous and the left/right limits of the derivative on the r.h.s. of (6.7) give the density in the coexisting phases. In our discussion we shall restrict ourselves to values of the parameters β and μ for which the system is in a unique phase. We can of course also go from the grand canonical pressure to the Helmholtz free energy density by the inverse of (6.5),

$$a(\beta, \rho) = \sup_{\mu} [\rho\mu - p(\beta, \mu)]. \tag{6.8}$$

Let $P(N_V \in \Delta|V| \,|\, \beta, \mu, V, b)$ be the probability of finding a total particle density in V (i.e., $N_V/|V|$) in the interval $\Delta := [n_1, n_2]$. Then, for b in the right class of b.c., we have (almost by definition) that

$$\lim_{V \nearrow \mathbb{R}^d} (\beta|V|)^{-1} \log P(N_V \in \Delta|V| \,|\, \beta, \mu, V, b) = \sup_{n \in \Delta} [a(\beta, \bar{\rho}) - a(\beta, n) + \mu(\bar{\rho} - n)], \tag{6.9}$$

were $\bar{\rho}$ is given by (6.7). In probabilistic language, this means that, up to a vertical translation, $-a(\beta, n) - \mu n$ is the LD functional, or rate function, for density fluctuations. (Note that $a(\beta, \rho)$ may be infinite for some values of ρ, i.e., when $\phi(r) = \infty$, for $r < D$, and ρ is above the close-packing density of balls with diameter D).

On the other hand, the fluctuations in all of V are clearly b.c. and ensemble dependent (they are non-existent in the canonical ensemble) and therefore not an intrinsic property of the system. Physically more relevant are the fluctuations not in the whole volume V but in a region Λ inside V. Of particular interest is the case when Λ is very large on the microscopic scale but still very small compared to V. The proper idealization of this situation is to first take the TL, $V \nearrow \mathbb{R}^d$, and then let Λ itself become very large. We are thus interested in the probability $P(N_\Lambda \in \Delta|\Lambda| \,|\, \beta, \mu)$, for Λ a large region in an infinite system obtained by taking the TL. This probability should now be an intrinsic property of a uniform single-phase macroscopic system characterized either by a chemical potential μ or by a density ρ.

A little thought shows that this probability corresponds to considering the grand canonical ensemble of a system of particles in a domain Λ with boundary interactions of the type

$$u_b(\mathbf{r}_i) = \sum_{k=1}^{\infty} \phi(|\mathbf{r}_i - \mathbf{x}_k|), \qquad \mathbf{r}_i \in \Lambda, \ \mathbf{x}_k \in \Lambda^c, \tag{6.10}$$

i.e., we imagine that the boundary interactions come from particles of the same type as those inside Λ, specified to be at positions $\mathbf{x}_1, \mathbf{x}_2, \ldots$ outside Λ. These positions must then be averaged over according to the infinite-volume Gibbs measure. It follows then, from the independence of the bulk properties of the system of the boundary conditions, that equation (6.9) is still correct, that is

$$\lim_{\Lambda \nearrow \mathbb{R}^d} (\beta|\Lambda|)^{-1} \log P(N_\Lambda \in \Delta|\Lambda| \mid \beta, \mu) = \sup_{n \in \Delta} [a(\beta, \bar{\rho}) - a(\beta, n) + \mu(\bar{\rho} - n)]. \tag{6.11}$$

This relation is indeed a theorem for classical systems, under fairly general conditions [G, vFS, O].

6.2 Quantum systems

It is equation (6.11) and similar formulas for fluctuations in the energy density which we want to generalize to quantum systems. To do this, we begin by considering the boundary conditions imposed on the N-particle wave functions $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N \mid V)$ for a quantum system in the domain V. Usually this is done by requiring that whenever any $\mathbf{r}_i$ is at the boundary of V, $\mathbf{r}_i \in \partial V$, then Ψ is equal to α times its normal derivative

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N \mid V) = \alpha\, \mathbf{n}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} \Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N \mid V) \tag{6.12}$$

with $\alpha = 0$ corresponding to Dirichlet and $\alpha = +\infty$ to Neumann boundary conditions.

Denote by b_α the elastic boundary condition (6.12). The existence of the TL of the grand canonical pressure $p(\beta, \mu \mid V, b_0)$ has been proven for quantum systems with stable potentials [R], and for positive potentials it is established that the pressure does not depend on α [Ro]. But, as far as we are aware, the dependence on $u_b(\mathbf{r}_i)$ has not been studied systematically, with the exception of the regime covered by the low-density expansion of Ginibre [Gi, BR]. This only shows that the dependence on the boundary is not so well understood for continuous quantum systems.

To investigate the density fluctuations in quantum systems we note that the momentum variables did not play any role in the derivation of (6.9) and (6.11) for classical systems. The only thing relevant, when considering particle number fluctuations, is the probability density in the configuration space. This is given for a classical system by integrating ν in (6.2) over the momentum variables, whose distribution is always a product of Gaussians (Maxwellians). For a quantum system, where the analog of (6.2) is the density matrix $\widehat{\nu}$, the configuration probability density is given by the diagonal elements of $\widehat{\nu}$ in the position representation. For the grand canonical ensemble this can be written as

$$\widehat{W}(\mathbf{R}_N \mid \beta, \mu, V, b_\alpha) := \frac{e^{\beta\mu N} \sum_\gamma |\Psi_\gamma(\mathbf{R}_N \mid V, b_\alpha)|^2 \, e^{-\beta E_\gamma}}{\widehat{\Xi}}, \tag{6.13}$$

where Ψ_γ and E_γ are the eigenstates and eigenvalues of H_N with the suitable statistics and b_α b.c. [R, B].

It is clear from the derivation of the TL [R, F] that, when $\phi(r)$ is super-stable, the TL for the canonical ensemble exists for all $\rho \in [n_1, n_2]$ with b.c. b_α. Then (6.9) carries over to quantum systems. The real problem is how to prove (6.11) for these systems. $\widehat{W}$ is no longer a Gibbs measure with a pair potential as interaction and there is no good reason to expect it to be a Gibbs measure for any other "reasonable" many-body potential [vFS]. (Even if the latter were the case, this potential would almost certainly depend on the density and temperature of the system and would therefore not carry directly any information on (6.11).) It might in fact appear that there is no strong reason why (6.11) should hold for quantum systems. The reason for expecting it to be true is that it is a thermodynamic type relation and such relations are in general unaffected by the transition from the classical to the quantum formalism. More explicitly, we see the difference between (6.9) and (6.11) as involving only boundary type quantities which should become irrelevant when Λ is of macroscopic size. The proof of such a statement seems far from obvious. In [LLS] this is achieved in the (technically) simplest case, where there are no interactions between the particles, i.e., the ideal gas with either Bose-Einstein or Fermi-Dirac statistics. We devote the rest of Part II to describe that work. In Chapter 7 we lay down the mathematical framework and state the main results, and in Chapter 8 we give the proofs. It turns out that even for these simple systems this project requires a certain amount of work. The more technical lemmas are proven in Appendix B.

To finish this introduction, we note that the same reasoning which leads to (6.11) also gives the well known result that, for $|\Lambda| \nearrow \infty$, the variance of N_Λ divided by $|\Lambda|$ tends to the compressibility, i.e., $\left\langle (N_\Lambda - \rho|\Lambda|)^2 \right\rangle / |\Lambda| \;\rightarrow\; \beta^{-1}(\partial^2 p/\partial\mu^2)(\beta, \mu)$. Furthermore the random variable $\xi := \lim_{|\Lambda| \nearrow \infty} (N_\Lambda - \rho|\Lambda|)/\sqrt{\rho|\Lambda|}$ satisfies a central limit theorem.

These results are also expected to remain valid for quantum systems and have been verified for certain classes of bosons [GVV, BMV, ABV], using algebraic methods. Here they are derived as a corollary of the LD result, and are presented in Section 8.1.

Chapter 7

Non-interacting quantum systems

We consider a d-dimensional square box $V := [-\ell/2, \ell/2]^d$. For computational convenience we choose periodic boundary conditions, but we do not expect our results to depend on this particular choice. (In fact we will restrict the thermodynamic parameters to the one-phase region.) In V there is an ideal fluid (either Fermi or Bose) in thermal equilibrium, as described by the grand canonical ensemble. We label the Bose fluid, shorthand BE, with the index $+$ and the Fermi fluid, shorthand FD, with the index $-$ and introduce the Fock space

$$\mathcal{F}_{\pm}^{V} := \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} L_{\pm}^{2}(V^n)\,, \tag{7.1}$$

where $L_{\pm}^{2}(V^n)$ is the n-particle space of all symmetric, resp. antisymmetric, square-integrable functions on V^n. Of course, for $n = 1$, $L_{\pm}^{2}(V) = L^2(V)$. In the sequel, in order to keep the notation light, we will often drop sub- or superscripts whenever there is no ambiguity.

Particles do not interact. Therefore the many-particle Hamiltonian in the box V can be written conveniently in the form

$$H_V := \bigoplus_{n=0}^{\infty} \sum_{i=1}^{n} \underbrace{1 \otimes \cdots \otimes h_V \otimes \cdots \otimes 1}_{i^{th} \text{ position out of } n}, \tag{7.2}$$

where h_V, the one-particle Hamiltonian on $L^2(V)$, is defined through the one-particle energy $\epsilon(k)$ in momentum space. This means that, if $|k\rangle$ denotes the momentum eigenvector (represented in $L^2(V)$ as $\psi_V^{(k)}(x) := e^{ik\cdot x}$), then $h_V|k\rangle = \epsilon(k)|k\rangle$ with $k \in V' := (2\pi\mathbb{Z}/\ell)^d$, the dual of V.

We assume $\epsilon(k)$ to be continuous, $\epsilon(0) = 0$ as a normalization, and $\epsilon(k) > 0$ for $k \neq 0$. Also $\epsilon(k) \approx |k|^\gamma$ for small k and $\epsilon(k) \geq |k|^\alpha$ for large k, with $\alpha, \gamma > 0$.

Furthermore we require

$$\int d^d x \left| \int d^d k \, e^{ik\cdot x} \frac{1}{e^{\beta\epsilon(k)-\beta\mu} - \varepsilon} \right| < \infty \tag{7.3}$$

for $\varepsilon = \pm 1$, $\beta > 0$, and suitable μ. (One might note the similarities between (7.3) and the space-clustering condition of [GVV, BMV], which ensures that generic observables there have normal fluctuations.)

The standard example of a non-relativistic, resp. relativistic, kinetic energy for a particle of mass m is $\epsilon(k) = k^2/(2m)$, resp. $\epsilon(k) = \sqrt{m^2c^4 + k^2c^2} - mc^2$ (having set Planck's constant $\hbar = 1$). Both functions satisfy the above conditions. The relativistic case includes $m = 0$, although this is not immediately obvious—cf. Appendix B.1 for details.

We observe that H_V may be rewritten as a quadratic form in the creation and annihilation operators on the Fock space $\mathcal{F}$. Let a_k^* be the operator that creates a particle in the state $|k\rangle$ and a_k the corresponding annihilator. Then

$$H_V = \sum_k \epsilon(k) a_k^* a_k = \sum_{j,k} \langle j|h_V|k\rangle a_j^* a_k =: \langle \mathbf{a}|h_V|\mathbf{a}\rangle \, . \tag{7.4}$$

We fix $\beta > 0$ and $\mu \in \mathbb{R}$ for FD, resp., $\mu < 0$ for BE. The grand canonical state in the volume V is defined by

$$\langle A \rangle_{\pm,\mu}^V := \frac{\mathrm{Tr}_{\mathcal{F}_\pm^V} \left(A \; e^{-\beta H_V + \beta\mu N} \right)}{\Xi_\pm^V(\mu)} \tag{7.5}$$

for every bounded operator A on $\mathcal{F}_\pm^V$. $N = N_V$ is the operator for the number of particles in the box V, $N|_{L^2_\pm(V^n)} := n 1_{L^2_\pm(V^n)}$, and $\Xi_\pm^V(\mu) := \mathrm{Tr}_{\mathcal{F}_\pm^V}(e^{-\beta H_V + \beta\mu N})$ denotes the partition function. As is well known (see, for example, [B]) we have

$$\Xi_+^V(\mu) = \prod_k \left(1 - e^{-\beta\epsilon(k)+\beta\mu} \right)^{-1} , \tag{7.6}$$

$$\Xi_-^V(\mu) = \prod_k \left(1 + e^{-\beta\epsilon(k)+\beta\mu} \right) . \tag{7.7}$$

The infinite-volume thermal state is defined through the limit

$$\langle \cdot \rangle := \lim_{V \nearrow \mathbb{R}^d} \langle \cdot \rangle^V , \tag{7.8}$$

when taking averages of local observables [BR, Sec. 2.6].

Taking the infinite volume limit of (7.6) and (7.7) one obtains the grand canonical pressure

$$p_\varepsilon(\mu) := \lim_{V\nearrow\mathbb{R}^d} \frac{\log \Xi_\varepsilon^V(\mu)}{\beta|V|} = -\frac{\varepsilon}{\beta(2\pi)^d}\int d^dk \log\left(1-\varepsilon\, e^{-\beta\epsilon(k)+\beta\mu}\right) \tag{7.9}$$

and the average density

$$\rho_\varepsilon(\mu) := \frac{dp_\varepsilon}{d\mu}(\mu) = \frac{1}{(2\pi)^d}\int d^dk\, \frac{1}{e^{\beta\epsilon(k)-\beta\mu}-\varepsilon}. \tag{7.10}$$

p_- is real analytic on the whole axis, whereas p_+ is real analytic only for $\mu < 0$ and has a finite limit as $\mu \to 0_-$. For convenience, we define $p_+(\mu) := \infty$ for $\mu > 0$. The slope of p_+ at 0_- is related to the Bose-Einstein condensation. We set

$$\rho_c := \rho(0_-) = \frac{1}{(2\pi)^d}\int d^dk\, \frac{1}{e^{\beta\epsilon(k)}-1}\,. \tag{7.11}$$

By the properties of $\epsilon(k)$, $\rho_c = \infty$ for $d \leq \gamma$, and is finite otherwise. ρ_c is the maximal density of the normal fluid and any surplus density is condensed into the $k = 0$ ground state. To simplify the notation we use ρ_c also in the case of an ideal Fermi fluid, setting it equal to ∞.

The infinite system is assumed to be in a pure thermal state, obtained through the limit (7.8) at the reference chemical potential μ. In this state the average density is $\bar\rho := \rho(\mu) < \rho_c$. We define the *translated pressure* by

$$g_{\varepsilon,\mu}(\lambda) = g_\varepsilon(\lambda) := p_\varepsilon(\mu+\lambda) - p_\varepsilon(\mu)\,. \tag{7.12}$$

g_ε is convex up, increasing, $g(0) = 0$, and $g'_\varepsilon(0) = \bar\rho$. For large negative values we have

$$\lim_{\lambda\to-\infty} g_\varepsilon(\lambda) = -p_\varepsilon(\mu)\,, \qquad \lim_{\lambda\to-\infty} g'_\varepsilon(\lambda) = 0, \tag{7.13}$$

whereas for positive values

$$\lim_{\lambda\to\infty} g_-(\lambda) = \infty\,, \qquad \lim_{\lambda\to\infty} g'_-(\lambda) = \infty \tag{7.14}$$

in the case of fermions and

$$\lim_{\lambda\to-\mu} g_+(\lambda) = p_+(0) - p_+(\mu)\,, \qquad \lim_{\lambda\to-\mu} g'_+(\lambda) = \rho_c \tag{7.15}$$

for bosons, with $g_+(\lambda) = \infty$ for $\lambda > -\mu$.

We define the *rate function* f_ε as the Legendre transform of g_ε, i.e.,

$$f_{\varepsilon,\mu}(x) = f_\varepsilon(x) := \inf_{\lambda\in\mathbb{R}} (g_\varepsilon(\lambda) - \lambda x) = g_\varepsilon(\lambda_o) - \lambda_o x. \tag{7.16}$$

Here $\lambda_o = \lambda_o(x)$ is the minimizer of $g(\lambda) - \lambda x$, which is unique by convexity. For $x \leq 0$ we have $\lambda_o = -\infty$. For $0 < x < \rho_c$, it is determined by $g'(\lambda_o) = x$, while for $x \geq \rho_c$ we have $\lambda_o = -\mu$. This shows that $f(x) = -\infty$ on the half-line $\{x < 0\}$ and finite elsewhere. In particular, f is convex down, strictly convex for $0 < x < \rho_c$, and $f_+(x) = p(0) - p(\mu) + \mu x$, for $x \geq \rho_c$, as a trace of the Bose-Einstein condensation.

Let us now consider a small subvolume Λ of our (already infinite) container V. The precise shape of Λ plays no role, only the "surface area" should be small compared to its volume $|\Lambda|$. Thus, by $\Lambda \nearrow \mathbb{R}^d$ we mean a sequence of subdomains such that for each Λ there exists a subset Λ' of Λ with $|\Lambda'|/|\Lambda| \to 1$ and $\mathrm{dist}(\Lambda', \mathbb{R}^d \setminus \Lambda) \to \infty$.

Let N_Λ be the number operator for the particles in Λ. With respect to $\langle\cdot\rangle$, N_Λ has some probability distribution. We follow the usual practice and use the same symbol N_Λ to denote also the corresponding random variable. Its distribution is indicated by $\mathbb{P}$, averages again by $\langle\cdot\rangle$.

We are now in a position to state the main result.

Theorem 7.1 *Let $\beta > 0$ and $\mu < 0$ for BE, resp. $\mu \in \mathbb{R}$ for FD. Then, for any interval $I := [a, b]$,*

$$\lim_{\Lambda\nearrow\mathbb{R}^d} \frac{1}{\beta|\Lambda|} \log \mathbb{P}(\{N_\Lambda \in |\Lambda| I\}) = \sup_{x\in I} f_{\varepsilon,\mu}(x).$$

Chapter 8

Large deviations in the density

This chapter is devoted to establishing Theorem 7.1. A way to gain the complete information about a random variable is to know its generating function. In the case at hand we want to obtain the *asymptotic* distribution of the "variable" N_Λ. Hence we need to study the asymptotic behavior of $\left\langle e^{\beta\lambda N_\Lambda}\right\rangle_\varepsilon$. This will be accomplished in Sections 8.2 and 8.3. In the case of bosons in high dimension, $\rho_c < \infty$ and one needs more detailed data to study the probability of supercritical densities. The necessary results are proven in Section 8.4.

For now, we start by showing how Theorem 7.1 follows from such an analysis via a standard LD technique. As a by-product, we also obtain a central limit theorem for $\xi_\Lambda := (N_\Lambda - \langle N_\Lambda\rangle)/|\Lambda|^{1/2}$.

8.1 The modified state

The following results are proved in the next sections.

Lemma 8.1 *There exists a $\lambda_{max}(\Lambda)$ such that $\left\langle e^{\beta\lambda N_\Lambda}\right\rangle_\varepsilon < \infty$ for all $\lambda < \lambda_{max}(\Lambda)$ and $\left\langle e^{\beta\lambda N_\Lambda}\right\rangle_\varepsilon = \infty$ for all $\lambda \geq \lambda_{max}(\Lambda)$. For FD we have $\lambda_{max}(\Lambda) = \infty$, whereas for BE $\lambda_{max}(\Lambda) < \infty$ with $\lambda_{max}(\Lambda) \searrow -\mu$, as $\Lambda \nearrow \mathbb{R}^d$.*

Theorem 8.2 *The limit*

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{\log\left\langle e^{\beta\lambda N_\Lambda}\right\rangle_\varepsilon}{\beta|\Lambda|} = g_\mu(\lambda)\,,$$

including any finite number of derivatives, exists uniformly on compacts of $\mathbb{R}$ for FD, resp. of $(-\infty, -\mu)$ for BE.

Inferring Theorem 7.1 from our information on the generating function is a standard argument from the theory of LD [E, O] (at least for subcritical densities $a < \rho_c$; the case $a \geq \rho_c$ requires more effort). Very roughly, the idea it to introduce a new state for which a certain event is typical.

Let us look at the statement of Theorem 7.1. The probability of the event in question can be rewritten as

$$Q_\Lambda := \left\langle \chi_{|\Lambda| I}(N_\Lambda) \right\rangle , \tag{8.1}$$

where χ_A is the indicator function of the set $A \subseteq \mathbb{R}$. To make this event typical we introduce the modified average

$$\langle \cdot \rangle_\lambda := \frac{1}{Z_\lambda} \left\langle \cdot \, e^{\beta \lambda N_\Lambda} \right\rangle , \tag{8.2}$$

where $\lambda < \lambda_{max}$ and the partition function $Z_\lambda := \left\langle e^{\beta \lambda N_\Lambda} \right\rangle$. With respect to this new state, (8.1) can be expressed as

$$Q_\Lambda = Z_\lambda \left\langle e^{-\beta \lambda N_\Lambda} \chi_{|\Lambda| I}(N_\Lambda) \right\rangle_\lambda . \tag{8.3}$$

The upper bound for Q_Λ comes from the exponential Chebychev inequality,

$$Q_\Lambda \leq \left\langle e^{\beta \lambda (N_\Lambda - a|\Lambda|)} \right\rangle = Z_\lambda \, e^{-\beta \lambda a |\Lambda|} \tag{8.4}$$

for any $0 < \lambda < \lambda_{max}$. As regards the lower bound we have to distinguish between two cases.

Case 1: $a < \rho_c$. One uses (7.13)-(7.15) to show that there exists a $\lambda_o < \lambda_{max}$ such that $g'(\lambda_o) = a$. Differentiating twice w.r.t. λ the equality in the statement of Theorem 8.2, we obtain

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{\langle N_\Lambda \rangle_{\lambda_o}}{|\Lambda|} = g'(\lambda_o) = \rho(\mu + \lambda_o) = a \, , \tag{8.5}$$

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{\beta}{|\Lambda|} \left[\langle N_\Lambda^2 \rangle_{\lambda_o} - \left(\langle N_\Lambda \rangle_{\lambda_o} \right)^2 \right] = \frac{d\rho}{d\mu}(\mu + \lambda_o) \, , \tag{8.6}$$

which is finite. This means that the event $\{N_\Lambda \approx |\Lambda| a\}$ is typical for the new state and a law of large numbers holds. Notice that $\lambda_o > 0$, since ρ is strictly increasing in μ.

From (8.3), $\forall c \in (a, b)$,

$$\begin{aligned} Q_\Lambda &\geq Z_{\lambda_o} \left\langle e^{-\beta\lambda_o N_\Lambda} \chi_{|\Lambda|[a,c]}(N_\Lambda) \right\rangle_{\lambda_o} \\ &\geq Z_{\lambda_o} \, e^{-\beta\lambda_o c|\Lambda|} \, \left\langle \chi_{|\Lambda|[a,c]}(N_\Lambda) \right\rangle_{\lambda_o} \\ &\geq \alpha \, Z_{\lambda_o} \, e^{-\beta\lambda_o c|\Lambda|} \end{aligned} \tag{8.7}$$

for some $\alpha \in (0, 1)$ and $|\Lambda|$ large. In fact, $\left\langle \chi_{|\Lambda|[a,c]}(N_\Lambda) \right\rangle_{\lambda_o} \to 1/2$, as $\Lambda \nearrow \mathbb{R}^d$. Therefore we obtain from (8.4), (8.7) and Theorem 8.2:

$$g(\lambda_o) - \lambda_o c + o(1) \leq \frac{\log Q_\Lambda}{\beta|\Lambda|} \leq g(\lambda_o) - \lambda_o a + o(1) \,. \tag{8.8}$$

Since $c \in (a, b)$ is arbitrary, we conclude that

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{\log Q_\Lambda}{\beta|\Lambda|} = g(\lambda_o) - \lambda_o a = f(a) = \sup_{x \in [a,b]} f(x) \,, \tag{8.9}$$

where f is the rate function defined in (7.16). λ_o is the same as in the definition of the Legendre transform (7.16), because of (8.5). The last equality comes from the convexity of f.

Case 2: $a \geq \rho_c$. In this case the problem is that one cannot find a *fixed* λ_o that verifies (8.5). As we will show later, it is nevertheless possible, for each finite Λ, to choose a λ_Λ such that the average density is a, i.e., $\langle N_\Lambda \rangle_{\lambda_\Lambda} = a|\Lambda|$. However, a might not correspond to the typical density in the limit $\Lambda \nearrow \mathbb{R}^d$, in the sense that no law of large numbers like (8.6) is guaranteed. Therefore establishing a lower bound for Q_Λ is not so immediate in this case, and we need the following lemma.

Lemma 8.3 *For every subdomain $\Lambda \subset \mathbb{R}^d$ and every $a > 0$, there exists a unique $\lambda_\Lambda = \lambda_\Lambda(a)$ such that $\langle N_\Lambda \rangle_{\lambda_\Lambda} = a|\Lambda|$. If $a \geq \rho_c$, then*

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \lambda_\Lambda = -\mu; \tag{8.10}$$

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{\log Z_{\lambda_\Lambda}}{\beta|\Lambda|} = g(-\mu); \tag{8.11}$$

$$\liminf_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{|\Lambda|} \ln \left\langle \chi_{[a|\Lambda|, a|\Lambda|+1)} \left(N_\Lambda\right) \right\rangle_{\lambda_\Lambda} = 0. \tag{8.12}$$

The above is proven in Section 8.4.

Now, the first two lines of (8.7) are still valid, with λ_Λ replacing λ_0. Taking the log and dividing by $|\Lambda|$, one obtains, via (8.12), the first inequality in (8.8), again for λ_Λ. The second inequality comes for free from Case 1. Finally, (8.10) and (8.11) are used to show that

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{\log Q_\Lambda}{\beta|\Lambda|} = g(-\mu) + \mu a = f(a) = \sup_{x \in [a,b]} f(x)\,, \tag{8.13}$$

yielding the linear part of the BE rate function. Q.E.D.

As anticipated, Theorem 8.2 also implies the central limit theorem for the density in Λ.

Corollary 8.4 *Under the assumptions of Theorem 7.1, the moments of the variable* $\xi_\Lambda = (N_\Lambda - \langle N_\Lambda \rangle)/|\Lambda|^{1/2}$ *converge, as* $\Lambda \nearrow \mathbb{R}^d$*, to those of a Gaussian with variance* $\beta^{-1}(d\rho/d\mu)(\mu)$*.*

Proof. The k-th cumulant of ξ_Λ is given by

$$C_\Lambda(k) = \frac{1}{\beta^k |\Lambda|^{k/2}} \left[\frac{d^k}{d\lambda^k} \log \left\langle e^{\beta\lambda N_\Lambda} \right\rangle \right]_{\lambda=0}\,, \tag{8.14}$$

$k \geq 2$. From Theorem 8.2, $C_\Lambda(2) \to \beta^{-1} g''(0) = \beta^{-1}(d\rho/d\mu)(\mu)$, whereas, for $k > 2$, $C_\Lambda(k) \to 0$. Also $C_\Lambda(1) = 0$. These limits are the cumulants of a centered Gaussian variable with the specified variance. Q.E.D.

8.2 Generating function

We derive a determinant formula for the generating function $\left\langle e^{\beta\lambda N_\Lambda} \right\rangle_\varepsilon$. With its help we prove the claims of Lemma 8.1. We will see in the next section that it is convenient to introduce the variables $\zeta := e^{\beta\lambda}$ and $\widetilde{\zeta} := \zeta - 1$.

By (7.4), we have

$$-\beta H_V + \beta\mu N_V = \langle \mathbf{a} | (-\beta h_V + \beta\mu 1_V) | \mathbf{a} \rangle \quad =: \quad \langle \mathbf{a} | A_V | \mathbf{a} \rangle, \tag{8.15}$$

$$\beta\lambda N_\Lambda = \langle \mathbf{a} | \beta\lambda\chi_\Lambda | \mathbf{a} \rangle \quad =: \quad \langle \mathbf{a} | B_\Lambda | \mathbf{a} \rangle, \tag{8.16}$$

which define A_V and B_Λ as linear operators on $L^2(V)$. We will use the following identity.

Lemma 8.5 *Let two linear operators A, B on $L^2(V)$ be self-adjoint and bounded from above. Then there exists a self-adjoint operator C on $L^2(V)$ such that $e^A e^B e^A = e^C$ and the relation*

$$e^{\langle \mathbf{a}|A|\mathbf{a}\rangle} e^{\langle \mathbf{a}|B|\mathbf{a}\rangle} e^{\langle \mathbf{a}|A|\mathbf{a}\rangle} = e^{\langle \mathbf{a}|C|\mathbf{a}\rangle}$$

holds for operators acting on $\mathcal{F}_+^V$ or $\mathcal{F}_-^V$.

Proof. See Appendix B.2.

We apply Lemma 8.5 with $A = A_V/2$ and $B = B_\Lambda$, after a symmetrization of the density matrix in (7.5). Then, using also definition (7.8),

$$\left\langle e^{\beta\lambda N_\Lambda} \right\rangle = \lim_{V\nearrow\mathbb{R}^d} \frac{\mathrm{Tr}_{\mathcal{F}^V}(e^{\langle \mathbf{a}|C|\mathbf{a}\rangle})}{\mathrm{Tr}_{\mathcal{F}^V}(e^{\langle \mathbf{a}|A_V|\mathbf{a}\rangle})}\,. \tag{8.17}$$

Evaluating the trace of a quadratic form in a_i^*, a_j is a standard calculation for both BE and FD. Let us consider first the case of fermions. For a self-adjoint operator A on $L^2(V)$ such that e^A is trace-class, we have

$$\mathrm{Tr}_{\mathcal{F}_-^V}(e^{\langle \mathbf{a}|A|\mathbf{a}\rangle}) = \det{}_V(1_V + e^A) = \det(1 + \chi_V e^A \chi_V), \tag{8.18}$$

where $\det_V$ is the determinant on $L^2(V)$ and det the determinant on $L^2(\mathbb{R}^d)$. Here and in the sequel we refer to the theory of infinite determinants, as found, e.g., in [RS, Sec. XIII.17]. e^{A_V} is obviously trace-class, and so is e^C, since e^{B_Λ} is bounded. Using the definition of C, we obtain

$$\begin{aligned}
\frac{\det_V(1_V + e^C)}{\det_V(1_V + e^{A_V})} &= \det{}_V\left[(1_V + e^{A_V})^{-1}(1_V + e^{A_V/2}e^{B_\Lambda}e^{A_V/2})\right] \\
&= \det{}_V\left[1_V + (1_V + e^{A_V})^{-1}e^{A_V/2}(e^{B_\Lambda} - 1_V)e^{A_V/2}\right] \\
&= \det\left[1 + \widetilde{\zeta}\,\chi_\Lambda\, D_{V,-}\,\chi_\Lambda\right],
\end{aligned} \tag{8.19}$$

where $D_{V,-} := (1+e^{A_V})^{-1}e^{A_V}$. We used the fact that $e^{B_\Lambda} = (e^{\beta\lambda}-1)\chi_\Lambda + 1 = \widetilde{\zeta}\chi_\Lambda + 1$ and the cyclicity of the trace in the definition of the determinant. Finally, from (8.17) and (8.19),

$$\left\langle e^{\beta\lambda N_\Lambda} \right\rangle_- = \lim_{V\nearrow\mathbb{R}^d} \det\left[1 + \widetilde{\zeta}\,\chi_\Lambda\, D_{V,-}\,\chi_\Lambda\right]. \tag{8.20}$$

One would like to take the limit on V inside the determinant by replacing $D_{V,-}$ with the corresponding operator on $L^2(\mathbb{R}^d)$ defined as

$$\widehat{(D_-\psi)}(k) := \widehat{d_-}(k)\widehat{\psi}(k), \qquad (D_-\psi)(x) = \int dy\, d_-(y-x)\,\psi(y), \tag{8.21}$$

where $\widehat{\ }$ denotes the Fourier transform and

$$\widehat{d_-}(k) := \frac{1}{1+e^{\beta(\epsilon(k)-\mu)}}. \tag{8.22}$$

Notice that $\widehat{d_-} \in L^1(\mathbb{R}^d)$ by our assumptions on $\epsilon(k)$ and so $d_- \in L^\infty(\mathbb{R}^d)$. Moreover, (7.3) ensures that $d_- \in L^1(\mathbb{R}^d)$.

By [RS, Sec. XIII.17, Lemma 4(d)] one has to establish that $\chi_\Lambda\, D_{V,-}\, \chi_\Lambda$ tends to $\chi_\Lambda\, D_-\, \chi_\Lambda$ in the trace norm.

Lemma 8.6 *Let $\widehat{d}$ be a continuous integrable function on $\mathbb{R}^d$. We define D through (8.21) as a linear operator acting on $L^2(\mathbb{R}^d)$. Furthermore we define D_V by $D_V|k\rangle := \widehat{d}(k)|k\rangle$ on $L^2(V)$ and by $D_V := 0$ on the orthogonal complement $L^2(\mathbb{R}^d \setminus V)$. Then, for $\Lambda \subset V$, $\chi_\Lambda D_V \chi_\Lambda$ and $\chi_\Lambda D \chi_\Lambda$ are trace-class, and*

$$\lim_{V \nearrow \mathbb{R}^d} \mathrm{Tr}|\chi_\Lambda(D_V - D)\chi_\Lambda| = 0.$$

Proof. See Appendix B.3.

We conclude that

$$\left\langle \zeta^{N_\Lambda} \right\rangle_- = \det(1 + \widetilde{\zeta}\,\chi_\Lambda\, D_-\, \chi_\Lambda), \tag{8.23}$$

with $\widetilde{\zeta} = \zeta - 1$.

For bosons we proceed in the same way, except that (8.18) is replaced by

$$\mathrm{Tr}_{\mathcal{F}_+^V}(e^{\langle \mathbf{a}|A|\mathbf{a}\rangle}) = \det{}_V(1_V - e^A)^{-1}, \tag{8.24}$$

requiring in addition $\|e^A\| < 1$. In fact, for $\|e^A\| \geq 1$, the l.h.s. of (8.24) is ∞, whereas the r.h.s. might be finite if 1 is not an eigenvalue of the trace-class operator e^A. In our case, by assumption $\|e^{A_V}\| < 1$. As for e^C, the function

$$\lambda \mapsto \|e^C\| = \|(e^{\beta\lambda} - 1)e^{A_V/2}\chi_\Lambda e^{A_V/2} + e^{A_V}\| \tag{8.25}$$

is increasing and $\lambda_{max}(\Lambda)$ is defined to be that λ which makes it equal to 1. Since the r.h.s. of (8.25) is increasing in Λ and its sup is $e^{\beta\lambda}\|e^{A_V}\| = e^{\beta(\lambda+\mu)}$, then one checks that $\lambda_{max}(\Lambda) \searrow -\mu$, as $\Lambda \nearrow \mathbb{R}^d$. Therefore, following the computation for FD, we have

$$\langle \zeta^{N_\Lambda} \rangle_+ = \lim_{V \nearrow \mathbb{R}^d} \frac{\det(1-e^C)^{-1}}{\det(1-e^{A_V})^{-1}} = \det(1+\widetilde{\zeta}\,\chi_\Lambda\, D_+\,\chi_\Lambda)^{-1}, \tag{8.26}$$

for $\lambda < \lambda_{max}$ and ∞ otherwise. Here D_+, the limit of $D_{V,+} := (e^{A_V}-1)^{-1}e^{A_V}$, is defined as in (8.21) with

$$\widehat{d_+}(k) := \frac{1}{1-e^{\beta(\epsilon(k)-\mu)}}. \tag{8.27}$$

Equation (8.26) is the analogue of (8.23) and proves Lemma 8.1.

8.3 Infinite volume limit

Instead of the chemical potential, in this section we use the fugacity $z := e^{\beta\mu}$, regarding it as a complex variable. This will come out handy for the proof of Theorem 8.2. The variables ζ and $\widetilde{\zeta}$, defined at the beginning of the previous section, will also be extended to the complex plane. In this setup the translated pressure (7.12) becomes

$$g_z(\zeta) = p(z\zeta) - p(z), \tag{8.28}$$

where, with a slight abuse of notation, we keep the same name for the pressure as a function of the fugacity.

Expressions (7.9)-(7.10) for the pressure and the average density define two analytic functions of μ in

$$E_+ \;:=\; \{\mathrm{Re}\,\mu < 0\} \cup \{\mathrm{Re}\,\mu \geq 0,\ \mathrm{Im}\,\mu \neq 2\pi j/\beta, \forall j \in \mathbb{Z}\}, \tag{8.29}$$

$$E_- \;:=\; \{\mathrm{Re}\,\mu < 0\} \cup \{\mathrm{Re}\,\mu \geq 0,\ \mathrm{Im}\,\mu \neq (2j+1)\pi/\beta, \forall j \in \mathbb{Z}\}. \tag{8.30}$$

Hence $g_\varepsilon(\zeta)$ is analytic in

$$G_+ := \mathbb{C} \setminus [z^{-1}, +\infty); \qquad G_- := \mathbb{C} \setminus (-\infty, -z^{-1}]. \tag{8.31}$$

We proceed to give the proof of Theorem 8.2. Let $K \subset G_\varepsilon$ be a compact set in the complex plane. We choose K such that $L := K \cap \mathbb{R}^+$ is also compact, since its image through the function $\zeta \mapsto \lambda$ verifies the hypotheses of the theorem. Our argument, however, is valid for any K. Without loss of generality, we can assume that $1 \in K$.

For $\widetilde{\zeta}$ restricted to $G_\varepsilon \cap \mathbb{R}^+$, let us define

$$\phi^\Lambda_{\varepsilon,z}(\zeta) := \frac{1}{|\Lambda|} \log \left\langle \zeta^{N_\Lambda} \right\rangle_{\varepsilon,z} = -\frac{\varepsilon}{|\Lambda|} \mathrm{Tr} \log(1 + \widetilde{\zeta}\chi_\Lambda D_\varepsilon \chi_\Lambda) \tag{8.32}$$

according to (8.23) and (8.26). The proof of Theorem 8.2 will be subdivided into three steps.

1. ϕ^Λ_ε can be analytically continued to G_ε.
2. There is a positive r such that $\phi^\Lambda(\zeta)$ converges uniformly to $\beta g_z(\zeta)$ for $|\zeta - 1| \le r$.
3. $|\phi^\Lambda|$ is uniformly bounded on K. Therefore by Vitali's Lemma [Ti, Sec. 5.21] $|\phi^\Lambda|$ and any finite number of its derivatives converge uniformly on K.

Step 1. We leave the proof of the following lemma to Appendix B.

Lemma 8.7 *The function $\phi^\Lambda_\varepsilon(\zeta)$, as defined by the trace in* (8.32)*, is analytic in G_ε.*

Step 2. Expanding the log in (8.32) one has, for $|\widetilde{\zeta}| < \|D_\varepsilon\|^{-1}$,

$$\phi^\Lambda_\varepsilon(\widetilde{\zeta} + 1) = -\frac{\varepsilon}{|\Lambda|} \mathrm{Tr} \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (\widetilde{\zeta}\chi_\Lambda D_\varepsilon \chi_\Lambda)^m. \tag{8.33}$$

We would like to interchange the summation with the trace. To do so, we need dominated convergence for the series:

$$|\widetilde{\zeta}\chi_\Lambda D_\varepsilon \chi_\Lambda|^m \le |\widetilde{\zeta}|^m \|D_\varepsilon\|^{m-1} (\chi_\Lambda D_\varepsilon \chi_\Lambda). \tag{8.34}$$

Since $|\Lambda|^{-1} \mathrm{Tr}(\chi_\Lambda D_\varepsilon \chi_\Lambda) = d_\varepsilon(0)$ (see proof of Lemma 8.6 in the Appendix), each term of (8.33) is bounded by a term of an integrable series independent of Λ. Therefore, for the same $\widetilde{\zeta}$'s as above,

$$\phi^\Lambda_\varepsilon(\widetilde{\zeta} + 1) = -\varepsilon \sum_{m=1}^{\infty} \frac{(-1)^{m-1}\widetilde{\zeta}^m}{m} \frac{1}{|\Lambda|} \mathrm{Tr}(\chi_\Lambda D_\varepsilon \chi_\Lambda)^m. \tag{8.35}$$

Suppose that we are able to prove that

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{|\Lambda|} \mathrm{Tr}(\chi_\Lambda D_\varepsilon \chi_\Lambda)^m = \int dk [\widehat{d_\varepsilon}(k)]^m, \tag{8.36}$$

with a rest bounded above by mR^m for some positive constant R. Then, using (8.22) and (8.27), we would have that, for any $r < \min\{\|D_\varepsilon\|^{-1}, R^{-1}\}$, uniformly for $|\widetilde{\zeta}| \le r$,

$$\begin{aligned}\lim_{\Lambda\nearrow\mathbb{R}^d} \phi_\varepsilon^\Lambda(\widetilde{\zeta}+1) &= -\varepsilon \sum_{m=1}^{\infty} \frac{(-1)^{m-1}\widetilde{\zeta}^m}{m} \int dk \left(\frac{1}{1-\varepsilon z^{-1} e^{\beta\epsilon(k)}}\right)^m \\ &= -\varepsilon \int dk \log\left(1 + \frac{\zeta - 1}{1-\varepsilon z^{-1} e^{\beta\epsilon(k)}}\right) \\ &= -\varepsilon \int dk \log\left(\frac{1-\varepsilon z\zeta e^{-\beta\epsilon(k)}}{1-\varepsilon z e^{-\beta\epsilon(k)}}\right) = \beta\, g_z(\widetilde{\zeta}+1),\end{aligned} \tag{8.37}$$

the last equality coming from (7.9). This would complete Step 2.

Let us pursue this project. One sees that

$$\begin{aligned}\int dk[\widehat{d}_\varepsilon(k)]^m &= \underbrace{(d_\varepsilon * d_\varepsilon * \cdots * d_\varepsilon)}_{m \text{ times}}(0) \\ &= \frac{1}{|\Lambda|}\int_\Lambda dx_1 \int_{\mathbb{R}^d} dx_2\, d_\varepsilon(x_1 - x_2) \cdots \int_{\mathbb{R}^d} dx_m\, d_\varepsilon(x_{m-1} - x_m)\, d_\varepsilon(x_m - x_1).\end{aligned} \tag{8.38}$$

The normalized integration over x_1 is harmless since, by translation invariance, the integrand does not depend on that variable. On the other hand, it is not hard to verify that

$$\begin{aligned}\frac{1}{|\Lambda|}\mathrm{Tr}(\chi_\Lambda D_\varepsilon \chi_\Lambda)^m &= \frac{1}{|\Lambda|}\int_\Lambda dx_1\, \langle x_1|(\chi_\Lambda D_\varepsilon \chi_\Lambda)^m|x_1\rangle \\ &= \frac{1}{|\Lambda|}\int_\Lambda dx_1 \int_\Lambda dx_2\, d_\varepsilon(x_1 - x_2) \cdots \int_\Lambda dx_m\, d_\varepsilon(x_{m-1} - x_m)\, d_\varepsilon(x_m - x_1).\end{aligned} \tag{8.39}$$

In view of (8.36), we want to compare (8.38) with (8.39). We observe that

$$\underbrace{\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}}_{m-1 \text{ times}} - \underbrace{\int_\Lambda\int_\Lambda\cdots\int_\Lambda}_{m-1 \text{ times}} = \sum_{i=1}^{m-1} \underbrace{\int_\Lambda\cdots\int_\Lambda}_{i-1 \text{ times}} \int_{\Lambda^c} \underbrace{\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}}_{m-1-i \text{ times}}. \tag{8.40}$$

Subtracting (8.38) from (8.39) leads then to $m-1$ terms of the form

$$\frac{1}{|\Lambda|}\int_\Lambda dx_1 \int_{\Lambda^c} dx_2\, d_\varepsilon(x_1 - x_2) \cdots \int_{A_m} dx_m\, d_\varepsilon(x_{m-1} - x_m)\, d_\varepsilon(x_m - x_1), \tag{8.41}$$

where the sets $A_3, \ldots, A_m$ can be either $\mathbb{R}^d$ or Λ. (8.41) holds because, due to the cyclicity of the integration variables, one can cyclically permute the order of integration without touching the integrand. We overestimate by switching to absolute values and integrating $x_3, \ldots, x_m$ over $\mathbb{R}^d$,

$$\frac{1}{|\Lambda|}\int_\Lambda dx_1\, (\chi_{\Lambda^c - x_1}|d_\varepsilon|) * |d_\varepsilon| * \cdots * |d_\varepsilon|)(x_1) =: \frac{1}{|\Lambda|}\int_\Lambda dx_1 u_\Lambda(x_1), \tag{8.42}$$

which defines $u_\Lambda(x_1)$. To estimate this function, we use recursively the relation $\|f * g\|_\infty \le \|f\|_\infty \|g\|_1$ and obtain

$$u_\Lambda(x_1) \le \|d_\varepsilon\|_1^{m-1} \sup_{\Lambda^c - x_1} |d_\varepsilon|. \tag{8.43}$$

Recalling now the definition of Λ' given before the statement of Theorem 7.1, one sees that, if $x_1 \in \Lambda'$ and $y \in \Lambda^c - x_1$, then $|y| \to \infty$ as $\Lambda \nearrow \mathbb{R}^d$. Hence, from (8.43),

$$\sup_{x_1 \in \Lambda'} \sup_{m \ge 1} \|d_\varepsilon\|_1^{-m+1} u_\Lambda(x_1) \to 0. \tag{8.44}$$

Also from (8.43), *pointwise in* x_1,

$$\sup_{m \ge 1} \|d_\varepsilon\|_1^{-m+1} u_\Lambda(x_1) \le \|d_\varepsilon\|_\infty. \tag{8.45}$$

When we average over $x_1 \in \Lambda$, the last two relations and the properties of Λ' prove that

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \sup_{m \ge 1} \|d_\varepsilon\|_1^{-m+1} \frac{1}{|\Lambda|} \int_\Lambda dx_1 u_\Lambda(x_1) = 0. \tag{8.46}$$

This takes care of each term as in (8.41), and we have $m-1$ of these terms. Hence (8.36) holds with $R = \|d_\varepsilon\|_1$. This ends Step 2.

Step 3. Again we expand (8.32) in powers of $\widetilde{\zeta}$, but this time about a generic $\widetilde{\zeta}_0 \notin G_\varepsilon - 1$ (see (8.31)). We obtain

$$\begin{aligned} \frac{1}{|\Lambda|} \mathrm{Tr} \log(1 + \widetilde{\zeta} \chi_\Lambda D_\varepsilon \chi_\Lambda) &= \frac{1}{|\Lambda|} \mathrm{Tr} \log(1 + \widetilde{\zeta}_0 \chi_\Lambda D_\varepsilon \chi_\Lambda) \\ + \quad \frac{1}{|\Lambda|} \mathrm{Tr} \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} &\Big((1 + \widetilde{\zeta}_0 \chi_\Lambda D_\varepsilon \chi_\Lambda)^{-1} \chi_\Lambda D_\varepsilon \chi_\Lambda \Big)^m (\widetilde{\zeta} - \widetilde{\zeta}_0)^m. \end{aligned} \tag{8.47}$$

Let us estimate this series. First of all, using some spectral theory [W, Sec. 7.4],

$$\begin{aligned} \|(1 + \widetilde{\zeta}_0 \chi_\Lambda D_\varepsilon \chi_\Lambda)^{-1}\| &\le \Big[\mathrm{dist}(1, \sigma(-\widetilde{\zeta}_0 \chi_\Lambda D_\varepsilon \chi_\Lambda)) \Big]^{-1} \\ &\le \Big[\mathrm{dist}(1, \sigma(-\widetilde{\zeta}_0 D_\varepsilon)) \Big]^{-1}, \end{aligned} \tag{8.48}$$

since we know from definitions (8.21), (8.22) and (8.27) that

$$\begin{aligned} \sigma(\chi_\Lambda D_- \chi_\Lambda) &\subset [0, \|\chi_\Lambda D_- \chi_\Lambda\|\,] \subset [0, \|D_-\|\,] \\ &= \sigma(D_-) = [0, 1/(1+z^{-1})], \end{aligned} \tag{8.49}$$

$$\begin{aligned} \sigma(\chi_\Lambda D_+ \chi_\Lambda) &\subset [-\|\chi_\Lambda D_+ \chi_\Lambda\|, 0] \subset [-\|D_+\|, 0] \\ &= \sigma(D_+) = [1/(1-z^{-1}), 0]. \end{aligned} \tag{8.50}$$

Repeating the same reasoning as in Step 2, we use the above to exchange the trace with the summation in (8.47)—which is legal for small $|\widetilde{\zeta}-\widetilde{\zeta}_0|$ as to be determined shortly. This yields a new series, whose m-th term is bounded above by

$$\Big[\mathrm{dist}(1,\sigma(-\widetilde{\zeta}_0 D_\varepsilon))\Big]^{-m}\|D_\varepsilon\|^{m-1}d_\varepsilon(0)\,|\widetilde{\zeta}-\widetilde{\zeta}_0|^m =: a\left(b(\widetilde{\zeta}_0)\,|\widetilde{\zeta}-\widetilde{\zeta}_0|\right)^m, \tag{8.51}$$

where $d_\varepsilon(0)=|\Lambda|^{-1}\mathrm{Tr}(\chi_\Lambda D_\varepsilon \chi_\Lambda)$. Hence, in view of (8.32), (8.47) implies

$$|\phi_\varepsilon^\Lambda(\widetilde{\zeta}+1)| \le |\phi_\varepsilon^\Lambda(\widetilde{\zeta}_0+1)| + a\frac{b(\widetilde{\zeta}_0)\,|\widetilde{\zeta}-\widetilde{\zeta}_0|}{1-b(\widetilde{\zeta}_0)\,|\widetilde{\zeta}-\widetilde{\zeta}_0|} \le |\phi_\varepsilon^\Lambda(\widetilde{\zeta}_0+1)|+a, \tag{8.52}$$

for $|\widetilde{\zeta}-\widetilde{\zeta}_0| \le (2b(\widetilde{\zeta}_0))^{-1}$.

The crucial fact is that $b(\widetilde{\zeta})^{-1}$ stays away from zero when $\widetilde{\zeta}$ is away from the boundary of $G_\varepsilon - 1$. This can be seen via the following argument, exploiting (8.51) and (8.49)-(8.50). In the FD case $\sigma(-\widetilde{\zeta}_0 D_-)$ is a segment that has one endpoint at the origin and the phase of $-\widetilde{\zeta}_0$ is the angle it forms with the positive semi-axis. This means that, as long as $\widetilde{\zeta}_0$ does not go anywhere near the negative semi-axis, we are safe. For $\widetilde{\zeta}_0 \in (-z^{-1}-1,0)$ (see (8.31)), $\sigma(-\widetilde{\zeta}_0 D_-)$ is contained in $\mathbb{R}_o^+$. However, notice from (8.49) that the other endpoint is located at $-\widetilde{\zeta}_0/(1+z^{-1})<1$. For BE the reasoning is analogous, except that in this case the phase of $\widetilde{\zeta}_0$ is the angle between $\sigma(-\widetilde{\zeta}_0 D_+)$ and $\mathbb{R}_o^+$. Therefore the "safe" span is the complement of the positive semi-axis. Also, if $\widetilde{\zeta}_0 \in (0, z^{-1}-1)$ (again see (8.31)), the "floating" endpoint of $\sigma(-\widetilde{\zeta}_0 D_+)$ is found at $\widetilde{\zeta}_0/(z^{-1}-1)<1$.

With the above estimate we can use (8.52) recursively. If $|\widetilde{\zeta}_0| \le r$, from Step 2, $|\phi_\varepsilon^\Lambda(\widetilde{\zeta}_0+1)| \le M$, for some M, since ϕ_ε^Λ converges uniformly there. Then, from (8.52), we have that $|\phi_\varepsilon^\Lambda(\widetilde{\zeta}_1+1)| \le M+a$, for any $\widetilde{\zeta}_1$ such that $|\widetilde{\zeta}_1-\widetilde{\zeta}_0| < (2b(\widetilde{\zeta}_0))^{-1}$. Proceeding, we see that $|\phi_\varepsilon^\Lambda(\widetilde{\zeta}_k+1)| \le M+ka$, whenever $|\widetilde{\zeta}_k-\widetilde{\zeta}_{k-1}| < (2b(\widetilde{\zeta}_{k-1}))^{-1}$. In this way we will cover K in finitely many steps since it keeps at a certain distance from the boundary of G_ε and the $(b(\widetilde{\zeta}_k))^{-1}$ are bounded below. This completes Step 3, i.e., $\phi_\varepsilon^\Lambda(\zeta)$ is bounded on K and Vitali's Lemma can be applied. Q.E.D.

8.4 Supercritical densities

For bosons in high dimension, Theorem 8.2 is not enough to establish the LD result for supercritical densities. In this section we prove Lemma 8.3. In particular, we derive a useful property of the distribution of N_Λ, w.r.t. the state $\langle \cdot \rangle_{\lambda_\Lambda}$, with λ_Λ chosen as stated in the lemma. $\varepsilon = +1$ will be understood in the reminder.

Letting the chemical potential go to zero in such a way that the average density remains constant (and bigger than ρ_c) is the usual way to proceed in the theory of Bose-Einstein condensation [B, ZUK, vLP1]. The limiting distribution of the global density $N_V/|V|$ is called *the Kac distribution*, and has been derived for several choices of V, $\epsilon(k)$ and d [ZUK, C, vLP1, vLL, LZP]. We do not go as far in this thesis. It is safe to say, however, that there is no reason to expect the distribution of $N_\Lambda/|\Lambda|$ to become degenerate (which would make (8.12) trivial).

PROOF OF LEMMA 8.3. Using a sloppy notation, let us write $\phi^\Lambda(\lambda)$ for $\phi^\Lambda(e^{\beta N_\Lambda})$ (see definition (8.32)). Differentiating this function, we get the mean density in the modified state: For $\lambda < \lambda_{max}$,

$$\rho^\Lambda(\lambda) := \frac{d\phi^\Lambda}{d\lambda}(\lambda) = \frac{1}{|\Lambda|} \frac{\langle N_\Lambda \, e^{\beta\lambda N_\Lambda} \rangle}{\langle e^{\beta\lambda N_\Lambda} \rangle} = \frac{\langle N_\Lambda \rangle_\lambda}{|\Lambda|}, \tag{8.53}$$

since, by virtue of Lemma 8.1, we can use dominated convergence to differentiate inside the average. Likewise,

$$v^\Lambda(\lambda) := \frac{d\rho^\Lambda}{d\lambda}(\lambda) = \frac{1}{|\Lambda|} \left[\langle N_\Lambda^2 \rangle_\lambda - (\langle N_\Lambda \rangle_\lambda)^2 \right]. \tag{8.54}$$

Thus, $\rho^\Lambda(\lambda)$ is increasing. The proof of Lemma 8.1 gives that $\lim_{\lambda \to \lambda_{max}} \rho^\Lambda(\lambda) = +\infty$. It is also clear that $\lim_{\lambda \to -\infty} \rho^\Lambda(\lambda) = 0$. The above implies the existence and uniqueness of λ_Λ such that $\rho^\Lambda(\lambda_\Lambda) = a$.

Let us fix $a \geq \rho_c$. Theorem 8.2 states in particular that, if $\Lambda \nearrow \mathbb{R}^d$, then $\rho^\Lambda(\lambda) \to \rho(\lambda) < \rho_c$, for $\lambda < -\mu$. This and Lemma 8.1 yield (8.10).

Continuing, let us choose some $\lambda_1 < -\mu$. From (8.10), $\lambda_1 \leq \lambda_\Lambda$, for Λ big enough. Using the monotonicity of ϕ^Λ and ρ^Λ,

$$\phi^\Lambda(\lambda_\Lambda) - \phi^\Lambda(\lambda_1) = \rho^\Lambda(\bar{\lambda}) \; (\lambda_\Lambda - \lambda_1) \leq a \; (\lambda_\Lambda - \lambda_1), \tag{8.55}$$

where $\bar{\lambda} \in (\lambda_1, \lambda_\Lambda)$ is given by Lagrange's Mean Value Theorem. Due to λ_1 being arbitrary, Theorem 8.2 and (8.55) imply

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \phi^\Lambda(\lambda_\Lambda) = \lim_{\lambda_1 \to -\mu} \beta\, g(\lambda_1) = \beta\, g(-\mu). \tag{8.56}$$

This is precisely (8.11).

The proof of (8.12) is more elaborate. Looking at (8.26) and the immediately following definition of $D_{V,+}$, we can write

$$\chi_\Lambda D_+ \chi_\Lambda =: (e^{-\beta(h'_\Lambda - \mu 1_\Lambda)} - 1_\Lambda)^{-1}\, e^{-\beta(h'_\Lambda - \mu 1_\Lambda)}, \tag{8.57}$$

which introduces a new one-particle Hamiltonian h'_Λ on $L^2(\Lambda)$. By Lemma 8.6, $\chi_\Lambda D \chi_\Lambda$ is trace-class, hence it has discrete spectrum. With the help of (8.50), we see that (8.57) can be solved for h'_Λ. Thus, h'_Λ is well defined and has the same spectral decomposition as $\chi_\Lambda D \chi_\Lambda$. In particular $\sigma(h'_\Lambda)$ is discrete. The eigenvalues of h'_Λ are indicated by $\epsilon'_j = \epsilon'_j(\Lambda)$, and are assumed to be in increasing order.

The idea behind this definition is to eliminate the two cut-offs χ_Λ in equation (8.26). These are responsible for all the complications in the proof of Theorem 8.2, and are the only manifestation that we are dealing with LD in Λ. (After all, the analogue of Theorem 8.2 for the global density $N_V/|V|$ is a trivial identity.) Definition (8.57) circumvents the problem in the sense that it incorporates the effect of the external volume V in the "local Hamiltonian" h'_Λ. One can then think of a system of free bosons in the container Λ and apply the available results for global fluctuations [ZUK, vLP2, LZP].

The drawback is that in general we have no precise information about $\sigma(h'_\Lambda)$. Even so, it is possible to determine the ground state of h'_Λ. In fact, by (8.57),

$$1_\Lambda + (e^{\beta\lambda} - 1_\Lambda)\chi_\Lambda D \chi_\Lambda = (1_\Lambda - e^{-\beta(h'_\Lambda - \mu 1_\Lambda)})^{-1}\,(1_\Lambda - e^{-\beta(h'_\Lambda - (\mu+\lambda) 1_\Lambda)}) \tag{8.58}$$

(cf. Step 2 in Section 8.3). Since $\chi_\Lambda D \chi_\Lambda$ is negative, the inf of the l.h.s. of (8.58) (in the sense of the quadratic form) is a decreasing function of λ, and attains zero at λ_{max}. This is so by the very definition of λ_{max}—see (8.25)-(8.26). On the other hand, the inf of the r.h.s. of (8.58) is zero if, and only if, $e^{-\beta(\epsilon'_0 - \mu - \lambda)} = 1$, whence

$$\epsilon'_0(\Lambda) = \lambda_{max}(\Lambda) + \mu > 0. \tag{8.59}$$

Let φ_j denote the eigenfunction relative to ϵ'_j, P_j the corresponding projector in $L^2(\Lambda)$, and a^*_j the creation operator on $\mathcal{F}^V$. We introduce $N_\Lambda^{(j)} := a^*_j a_j$, the operator for the number of particles in the state φ_j. $\{N_\Lambda^{(j)}\}$ is a commuting family and $N_\Lambda = \sum_{j=0}^\infty N_\Lambda^{(j)}$. We want to verify that these operators behave like independent random variables w.r.t. $\langle \cdot \rangle_\lambda$. We can study their joint generating function, employing the same techniques as in Section 8.2. In fact, for η_j bounded, define $B'_\Lambda := \sum_j \eta_j P_j + \beta\lambda\chi_\Lambda$ and replace (8.16) with

$$\langle \mathbf{a}|B'_\Lambda|\mathbf{a}\rangle = \sum_{j=0}^\infty \eta_j N_\Lambda^{(j)} + \beta\lambda N_\Lambda. \tag{8.60}$$

One verifies that, in $L^2(V)$ or in $L^2(\mathbb{R}^d)$, $e^{B'_\Lambda} = \sum_j e^{\eta_j+\beta\lambda}P_j + 1 - \chi_\Lambda$. In particular $e^{B'_\Lambda} - 1 = \chi_\Lambda(e^{B'_\Lambda} - 1)\chi_\Lambda$. This allows us to proceed as in (8.26), and write

$$Z_\lambda \left\langle e^{\sum_j \eta_j N_\Lambda^{(j)}} \right\rangle_\lambda = \left\langle e^{\langle \mathbf{a}|B'_\Lambda|\mathbf{a}\rangle} \right\rangle = \lim_{V\nearrow\mathbb{R}^d} \det\left[1 + (e^{B'_\Lambda} - 1)\chi_\Lambda D_V \chi_\Lambda\right]^{-1}. \tag{8.61}$$

Taking the above limit is slightly more complicated than the corresponding computation in Section 8.2. Since $e^{B'_\Lambda} - 1$ is bounded,

$$\mathrm{Tr}|(e^{B'_\Lambda} - 1)\chi_\Lambda(D_V - D)\chi_\Lambda| \le \|e^{B'_\Lambda} - 1\|\ \mathrm{Tr}|\chi_\Lambda(D_V - D)\chi_\Lambda|. \tag{8.62}$$

One then applies Lemma 8.6 and [RS, Sec. XIII.17, Lemma 4(d)], so that (8.61) gives

$$\begin{aligned}\left\langle e^{\sum_j \eta_j N_\Lambda^{(j)}} \right\rangle_\lambda &= Z_\lambda^{-1} \det\left[1 + (e^{B'_\Lambda} - 1)\chi_\Lambda D \chi_\Lambda\right]^{-1} \\ &= Z_\lambda^{-1} \det\nolimits_\Lambda \left[1_\Lambda + (e^{B'_\Lambda} - 1_\Lambda)(e^{-\beta(h'_\Lambda - \mu 1_\Lambda)} - 1_\Lambda)^{-1} e^{-\beta(h'_\Lambda - \mu 1_\Lambda)}\right]^{-1} \\ &= \prod_{j=0}^\infty \frac{1 - e^{-\beta(\epsilon'_j - \mu - \lambda)}}{1 - e^{-\beta(\epsilon'_j - \mu - \lambda) + \eta_j}},\end{aligned} \tag{8.63}$$

having used (8.26) and (8.58) to express Z_λ.

This shows that the $N_\Lambda^{(j)}$'s represent a set of independent, geometrically distributed random variables, with averages $\langle N_\Lambda^{(j)}\rangle_\lambda = (e^{\beta(\epsilon'_j - \mu - \lambda)} - 1)^{-1}$. At this point we can apply [vLP2, Lemma 2] to $N_\Lambda^{(0)}$ and $N_\Lambda - N_\Lambda^{(0)}$. We obtain

$$\left\langle \chi_{[a|\Lambda|, a|\Lambda|+1)}(N_\Lambda)\right\rangle_{\lambda_\Lambda} \ge \frac{1}{a|\Lambda|}\, e^{-\beta(\epsilon'_0(\Lambda) - \mu - \lambda_\Lambda)(a|\Lambda|+2)}. \tag{8.64}$$

Assertion (8.12) is derived from (8.64) via (8.10), (8.59) and Lemma 8.1. Q.E.D.

Part III

Appendices

Appendix A

Escape orbits for smooth non-compact billiards

Let f be a smooth function from $\mathbb{R}_o^+$ to $\mathbb{R}^+$, bounded, vanishing when $x \to +\infty$. No integrability assumptions are given. Now consider the billiard flow in the table $\Omega := \{(x, y) \in \mathbb{R}_o^+ \times \mathbb{R}_o^+ \mid 0 \le y \le f(x)\}$, as shown in Fig. A.1.

The analogies between these types of non-compact billiards and the infinite step billiards of Part I are evident. So it is sound to ask ourselves similar questions: in particular, how many escape orbits can we have in Ω? (In this case the obvious notion of escape orbit, which we give nonetheless for the sake of completeness, corresponds to a trajectory that has a positive x-velocity for—say—all positive times.) Of course there is always one of these: it is the orbit that lies on the horizontal semi-axis. We call it the *trivial orbit*.

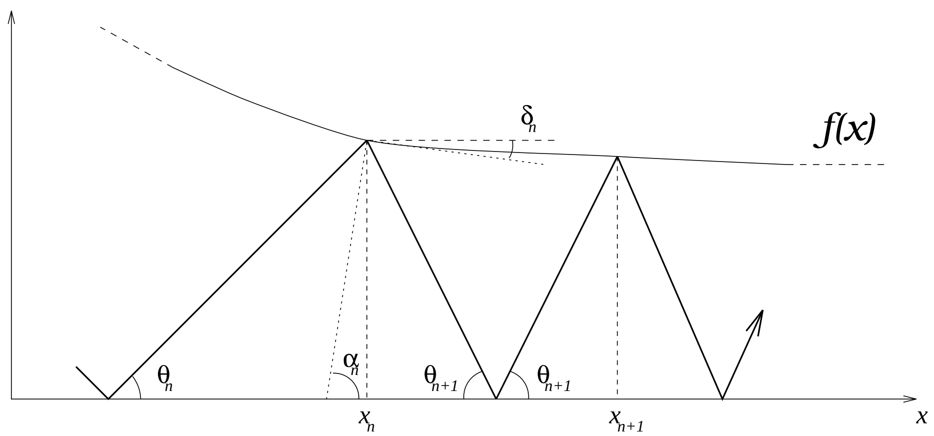

Figure A.1: Sketch of an escape orbit in the billiard Ω.

We report here the conclusions of [Le], which shows that, in great generality, there is no other escape orbit than the trivial one.

Apparently, the question was first touched on in [L, Thm. 2], where the above result is obtained for *eventually convex* f's, i.e., convex in a neighborhood of ∞. The same statement is found in the more recent review article [Ki]. It can be explained easily, at least for tables with finite area. In that case the cusp has an asymptotically vanishing measure. If we had an escape orbit, then, due to the hyperbolicity of the flow in that region of the phase space, we could also find a non-zero measure set of escape orbits—which fall into the cusp—with an obvious contradiction. (This argument is described in [Ki] as “A gallon of water won't fit inside a pint-sized cusp”.) In fact, let us suppose there is an escape orbit, whose initial point we can always fix at $(x_0, y_0) = (x_0, f(x_0))$, on the upper boundary of Ω, far enough to lie in the region where f is convex. The initial velocity has an x-component $v_x > 0$. Now it is easily seen that every other set of initial conditions $x_0' \geq x_0$ and $v_x' \geq v_x$ (provided v and v' have the same modulus) leads to a new escape orbit, due to the dispersing effect of the convex upper wall.

The same argument may be used to deduce that an infinite cusp on the Poincaré's disc does not allow non-trivial trajectories to collapse into it, which is implicitly stated in [GiU].

We are going to relax the hypothesis for the non-existence result to hold: asymptotic hyperbolicity is not a necessary condition at all. We may allow f to have flex points and abrupt negative variations (see, for instance, Fig. A.2,(c)). The proofs are presented in the next section, while examples (and counter-examples) are discussed in Section A.2.

A.1 The result

Theorem A.1 *Consider the plane billiard table generated as above by the function f defined on $\mathbb{R}_o^+$, twice differentiable, positive, bounded, such that*

$$f(x) \searrow 0 \text{ as } x \to +\infty. \tag{A.1}$$

Also, for x sufficiently large,

$$f'(x) < 0, \tag{A.2}$$

Then, under either one of these assumptions:

$$\limsup \frac{f'(x)}{f(x)} < 0; \tag{A.3}$$

or

$$\lim_{x\to+\infty} f'(x) = 0 \qquad \textit{and} \qquad \limsup_{x\to+\infty} \frac{f''f}{f'}(x) < +\infty; \tag{A.4}$$

no orbits but the trivial one are escape orbits.

Remark. The assumption about the convexity of f is contained in (A.4): in fact, if $f'' \geq 0$ then necessarily $f' \nearrow 0$ and $f''f/f' \leq 0$.

Proof of Theorem A.1. Suppose that we have a non-trivial escape orbit. Let us fix, without loss of generality, an initial point in a neighborhood of $+\infty$ where all the asymptotic hypotheses hold. For instance, (A.2) would do, and (A.4), if this is the case, would be read as

$$f'(x) \geq -\varepsilon \qquad \text{and} \qquad \frac{f''f}{f'}(x) < k_1, \tag{A.5}$$

for some $\varepsilon > 0$. Also, for some consistency in the notation, let us suppose the initial point to be a bouncing point on the upper wall, i.e., $(x_0, y_0) = (x_0, f(x_0))$.

Using the notations of Fig. A.1 we have the fundamental relation:

$$\tan\theta_{n+1}(x_{n+1} - x_n) = f(x_n) + f(x_{n+1}). \tag{A.6}$$

With a bit of geometry, looking at the same picture, we get

$$\theta_{n+1} = \theta_n + \pi - 2\alpha_n = \theta_n + 2\delta_n, \tag{A.7}$$

where $\delta_n := -\arctan(f'(x_n)) > 0$. This summarizes to

$$\theta_n = \theta_1 + 2\sum_{k=1}^{n-1} \delta_k. \tag{A.8}$$

Thus, $\{\theta_n\}_{n\geq 1}$ is an increasing sequence. Since we have assumed the particle to never travel backwards, then $\theta_n < \pi/2$ for all $n \geq 1$, so $\theta_n \nearrow \theta_\infty \in [\theta_1, \pi/2]$. Hence $\tan\theta_n \geq \tan\theta_1 =: k_2 > 0$. From this, the monotonicity of f, and (A.6) we have

$$x_{n+1} - x_n \leq \frac{2}{k_2} f(x_n). \tag{A.9}$$

What is stated above implies that $\sum_k \delta_k < +\infty$. Therefore $\delta_k \to 0$. As a consequence, we see that $\exists k_3 \in]0,1[$ such that $\delta_k \geq k_3 \tan \delta_k = k_3 |f'(x_k)|$. If we can prove that

$$-\sum_{k=0}^{\infty} f'(x_k) = +\infty, \tag{A.10}$$

then $\sum_k \delta_k$ cannot converge, creating a contradiction which finishes the proof.

Defining $g := -f'/f > 0$ will greatly simplify our notation. From (A.9) we obtain, for some constant k_4,

$$-\sum_n f'(x_n) \geq k_4 \sum_n g(x_n)(x_{n+1} - x_n). \tag{A.11}$$

Case (A.3). Obviously (A.3) reads $g \geq k_5 > 0$. Hence (A.11) gives (A.10), since by hypothesis $x_n \to \infty$. It may be worth recalling that (A.3) means we have exponential decay for f. In fact, after an integration, we see that $\forall x_2 > x_1 \geq 0$,

$$f(x_2) \leq f(x_1)\, e^{-k_5(x_2 - x_1)}. \tag{A.12}$$

Case (A.4). The proof is a little more involved here. We use our assumption on the limit of f' to prove the following.

Lemma A.2 *There exists a constant k_6 such that $\forall n \in \mathbb{N}$*

$$\frac{f(x_n)}{f(x_{n+1})} \leq k_6.$$

Proof. Let us call $\bar{x}_n$ the point provided by Lagrange's Mean Value Theorem applied to f in $[x_n, x_{n+1}]$. We can write

$$\frac{f(x_n)}{f(x_{n+1})} = 1 - \frac{f'(\bar{x}_n)(x_{n+1} - x_n)}{f(x_{n+1})}. \tag{A.13}$$

Using (A.9), this is turned into

$$\frac{f(x_n)}{f(x_{n+1})} \left(1 + \frac{2f'(\bar{x}_n)}{k_2}\right) \leq 1, \tag{A.14}$$

which yields the lemma, since the term in the brackets is positive for n large enough, because of the assumption about the vanishing of f'. Q.E.D.

This enables us to derive a further lemma.

Lemma A.3 *There exists a constant* $k_7 \in]0,1[$ *such that, for* n *sufficiently large,* $g(x_n) \geq k_7 \max_{[x_n,x_{n+1}]} g$.

PROOF. Proving the statement is equivalent to finding a $k_8 > 0$ for which

$$\log g(\widetilde{x}_n) - \log g(x_n) \leq k_8, \tag{A.15}$$

where $\widetilde{x}_n$ maximizes g in $[x_n, x_{n+1}]$. Using again Lagrange's Mean Value Theorem, the fact that $\widetilde{x}_n - x_n \leq (2/k_2) f(x_n)$—a consequence of (A.9)—and the previous lemma, we obtain

$$\begin{aligned} \log g(\widetilde{x}_n) - \log g(x_n) &= \frac{g'}{g}(\widehat{x}_n)(\widetilde{x}_n - x_n) \\ &\leq \frac{2}{k_2}\left(\frac{f''}{f'} - \frac{f'}{f}\right)(\widehat{x}_n)\ f(x_n) \\ &\leq k_9\left(\frac{f''}{f'} - \frac{f'}{f}\right)(\widehat{x}_n)\ f(x_{n+1}) \\ &\leq k_9\left(\frac{f''f}{f'} - f'\right)(\widehat{x}_n) \leq k_9(k_1 + \varepsilon), \end{aligned} \tag{A.16}$$

having used (A.5) in the last step. Q.E.D.

We are now prompted to get (A.10) in this case, too. Looking at (A.11) we can write:

$$\begin{aligned} \sum_n g(x_n)(x_{n+1} - x_n) &\geq k_7 \sum_n (\max_{[x_n,x_{n+1}]} g)(x_{n+1} - x_n) \\ &\geq k_7 \int_{x_0}^{+\infty} g(x)dx = +\infty, \end{aligned} \tag{A.17}$$

since $-\int^{\infty}(f'/f) = -\lim_{x\to\infty}(\log f(x) + const) = +\infty$. Q.E.D.

A.2 Discussion

The evident news the theorem puts forth, compared to the mentioned condition $f'' > 0$, is the possibility for f' to oscillate, to a certain extent. Dynamically speaking, the change in direction our particle gets every time it bounces against the upper wall ($\delta_k = -\arctan(f'(x_k))$, that is) need not be a monotonic sequence. As a matter of fact, (A.4) precisely controls the amount of such an oscillation. An example will illustrate the case: for $\alpha > 1, \beta > 0, c > 1$ define $f_1'(x) := -x^{-\alpha}(\sin(x^\beta) + c) < 0$. This means that we define $f_1(x) := -\int_x^\infty f_1'(z)dz$, which makes sense as a convergent integral. Therefore $f_1''(x) = -\beta x^{-\alpha+\beta-1}\cos(x^\beta) + O(x^{-\alpha-1})$, showing that f_1 is not convex. Now the asymptotic behavior of f_1 and its derivatives is easily derived:

$$\frac{f_1'' f_1}{f_1'}(x) \approx x^{-\alpha+\beta}. \tag{A.18}$$

Thus, (A.4) holds if, and only if, $\alpha \geq \beta$, meaning that the faster f_1 vanishes the more violent the oscillation of f_1' is allowed to be.

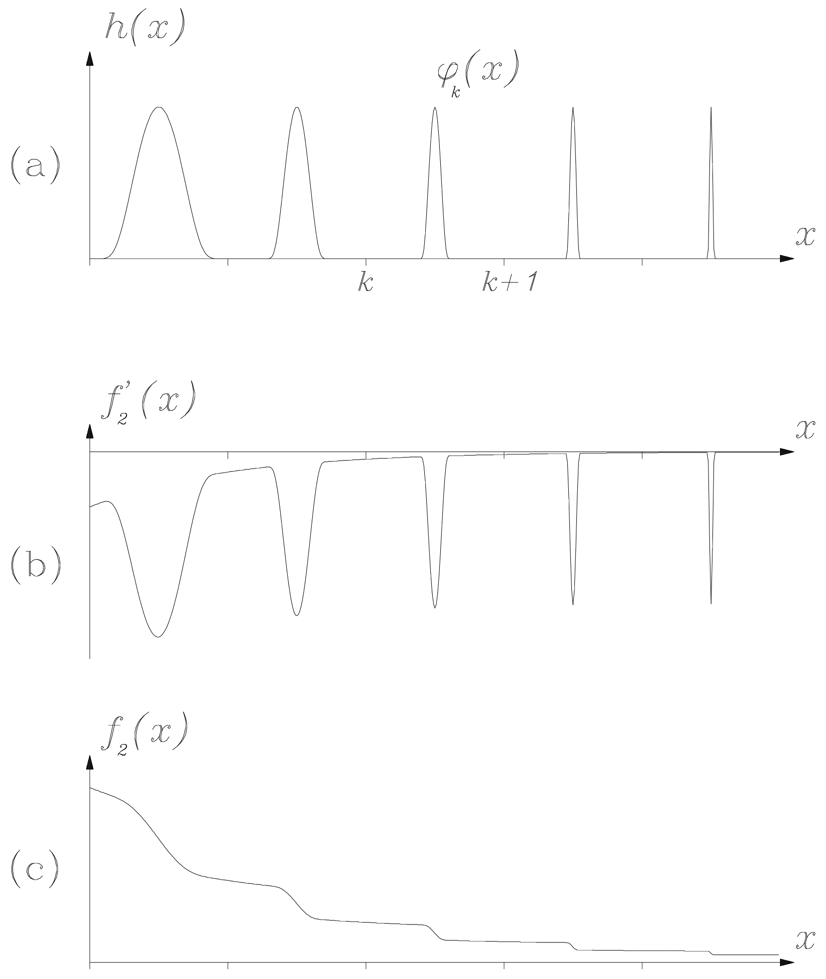

Figure A.2: Construction of a billiard table satisfying (A.3) but not (A.4).

Another example may be interesting to present, to show that there are cases where (A.3) holds but (A.4) does not. Pick a $\phi \in C^\infty(\mathbb{R})$ supported in $]-1/2, 1/2[$, with

$\int \phi = 1 - e^{-1}$. Then call, for $k \in \mathbb{N}$,

$$\phi_k(x) := \phi((x - k - 1/2)\, e^k), \tag{A.19}$$

which is therefore supported in $]k + 1/2 - e^{-k}/2, k + 1/2 + e^{-k}/2[$. Let us define $h(x) := \sum_{k=0}^{\infty} \phi_k(x)$. The result is shown in Fig. A.2,(a). We see that

$$\int_k^{k+1} h(x)dx = \int_k^{k+1} \phi_k(x)dx = (1 - e^{-1})e^{-k}. \tag{A.20}$$

Also denote by $H(x) := \int_x^{\infty} h(z)dz$. Finally, let us introduce $f_2'(x) := -e^{-x} - h(x)$, corresponding to $f_2(x) = e^{-x} + H(x)$. Their graphs are displayed in Fig. A.2, (b) and (c), respectively. Certainly $f_2' \not\to 0$ and (A.4) is not verified. We can estimate f_2 from above. In fact $e^{-[x]-1} \leq H(x) \leq e^{-[x]}$, giving $H(x) \leq e^{-x+1}$. Therefore

$$\frac{|f_2'|}{f_2}(x) = \frac{e^{-x} + h(x)}{e^{-x} + H(x)} \geq \frac{e^{-x}}{e^{-x} + e^{-x+1}} = \frac{1}{1+e} \geq 0. \tag{A.21}$$

That is: (A.3) holds as well as the result, in this case.

Is it difficult to say to what extent our theorem is inclusive of the general case or how it can be refined. The point here is that finding a sufficient condition for the non-existence of an escape orbit is much more direct than finding a necessary condition. The shape of f can be *pathological* enough but not in a suitable way that allows a trajectory to go directly to infinity. Of one thing we can be assured, though: hypotheses (A.1)-(A.2) do not suffice and one needs some extra assumption to control a possible wild behavior of f'. To show this point, we proceed to construct a billiard table verifying those hypotheses and having one escape orbit. We will start by first drawing the orbit and then a compatible f.

Consider the polyline shown in Fig. A.3,(a) with $\theta_1 \in]0, \pi/2[$ and $\{y_n\}$, any non-integrable sequence such that $y_n \searrow 0$. If this were an escape orbit then we would have $f(x_n) = y_n$, $f'(x_n) = 0$ and $\theta_n = \theta_1\ \forall n$. Furthermore $x_{n+1} - x_n = (y_{n+1} + y_n)/ \tan\theta_1$ so that $\lim_{n\to\infty} x_n = \infty$, because of our assumption on $\{y_n\}$. Of course any f giving rise to such an orbit cannot satisfy (A.2), because of the flat tangent at the bouncing points, but we can slightly modify our picture in order to fit it. Take an integrable

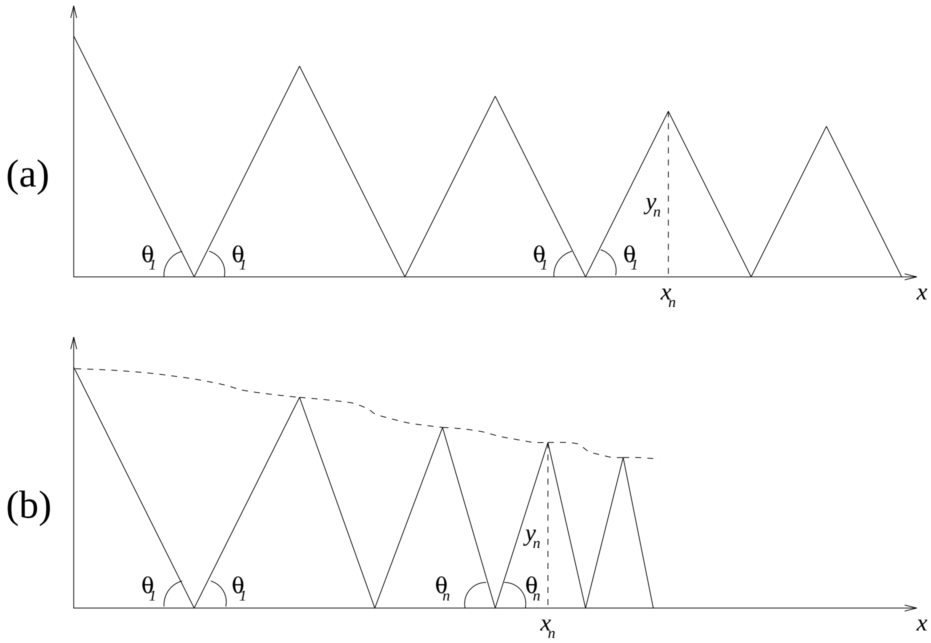

Figure A.3: Construction of a billiard, verifying neither (A.3) nor (A.4), in which a non-trivial escape orbit os found.

sequence $\{\delta_n\}$, $\delta_n > 0$ such that $\theta_\infty := \theta_1 + 2\sum_n \delta_n < \pi/2$. Now modify the trajectory in Fig. A.3,(a), "shrinking" it in order to have $\theta_n := \theta_1 + 2\sum_{k=1}^{n-1} \delta_k$; keep y_n fixed. The result is drawn in Fig. A.3,(b). This is again an escape orbit since, due to our choice of θ_∞, the contraction of the little triangles has a lower bound, i.e. $x_{n+1} - x_n \geq (y_{n+1} + y_n)/\tan\theta_\infty$. One can now very easily construct an f which satisfies (A.1) and (A.2) and whose graph is an upper wall for this trajectory.

This proves our remark.

Appendix B

Technical lemmas for Part II

We collect here the proofs of the more technical lemmas of Part II.

B.1 Relativistic massless particles

We prove that the energy dispersion $\epsilon(k) = c|k|$ satisfies our assumptions. The only condition to be checked is (7.3), that is, the Fourier transform of $k \mapsto (e^{\beta(c|k|-\mu)} - \varepsilon)^{-1}$ is in $L^1(\mathbb{R}^d)$. This is a consequence of the following

Lemma B.1 *Let $f : [0, +\infty) \longrightarrow \mathbb{C}$ be of Schwartz class. With the common abuse of notation, denote by $\widehat{f}(|\xi|)$ the Fourier transform of $f(|x|)$, for $x, \xi \in \mathbb{R}^d$. Then, for some positive C,*

$$\widehat{f}(|\xi|) \leq \frac{C}{|\xi|^{d+1}}.$$

Proof. For simplicity let us write $\xi = |\xi|$. The Fourier transform of a radial function is

$$\widehat{f}(\xi) = \frac{(2\pi)^{d/2}}{\xi^{d/2-1}} \int_0^\infty dr\, f(r)\, r^{d/2} J_{d/2-1}(r\xi)\,, \tag{B.1}$$

cf. [SW, Chap. IV, Thm. 3.3], where J_ν is the standard Bessel function of order ν [Wa]. One has

$$J_\nu(x) \approx \frac{x^\nu}{2^\nu \Gamma(\nu+1)}, \tag{B.2}$$

for $x \to 0$, whereas

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}} \left[\cos\left(x - \frac{2\nu+1}{4}\,\pi \right) + g(x) \right], \tag{B.3}$$

with $g(x) \to 0$ for $x \to \infty$. Using the relation

$$\int_0^x dt\, t^\nu J_{\nu-1}(t) = x^\nu J_\nu(x) \tag{B.4}$$

we integrate (B.1) by parts repeatedly, taking into account also (B.2) and the hypothesis on f. After n integrations we get, up to constants, n terms of the form

$$\frac{1}{\xi^{d/2+n-1}} \int_0^\infty dr\, f^{(i)}(r)\, r^{d/2-n+i} J_{d/2+n-1}(r\xi), \tag{B.5}$$

with $i = 1, \ldots, n$. For our purposes it suffices to iterate up to $n \geq d/2 + 2$. In fact, if i is such that $d/2 - n + i > -d$, then in (B.5) we can estimate the Bessel function by a constant. The integral converges by the rapid decay of $f^{(i)}$ and the whole term is of the order ξ^{-d-1} or better. For smaller values of i, the estimate uses (B.2), for $x \in [0, a]$, and (B.3) otherwise. Since $|f^{(i)}| \leq c$, (B.5) is bounded by

$$\begin{aligned} & \frac{A}{\xi^{d/2+n-1}} \int_0^{a/\xi} dr\, r^{d/2-n+i} (r\xi)^{d/2+n-1} + \\ + \; & \frac{B}{\xi^{d/2+n-1}} \int_{a/\xi}^\infty dr\, r^{d/2-n+i} (r\xi)^{-1/2} \approx \frac{1}{\xi^{d+i}}, \end{aligned} \tag{B.6}$$

the second integral being convergent because of the choice of i. Q.E.D.

B.2 Proof of Lemma 8.5

As before, we set $\varepsilon = \pm 1$, according to either bosons or fermions. A general $f \in L^2(V)$ can be expanded in the Fourier basis as $f = \sum_k f_k |k\rangle$. The corresponding creation operator is then defined by

$$a(f)^* := \sum_{k \in V'} f_k\, a_k^*. \tag{B.7}$$

For the sake of simplicity, we denote $\mathcal{A} := \langle \mathbf{a} | A | \mathbf{a} \rangle = \sum_{ij} A_{ij} a_i^* a_j$ (same for $\mathcal{B}$). Recalling the canonical (anti)commutation relations,

$$\begin{aligned} [a_i, a_j^*]_{-\varepsilon} &= a_i a_j^* - \varepsilon\, a_j^* a_i = \delta_{ij}; \\ [a_i, a_j]_{-\varepsilon} &= a_i a_j - \varepsilon\, a_j a_i = 0, \end{aligned} \tag{B.8}$$

one calculates that

$$[\mathcal{A}, a(f)^*] = a(Af)^*, \tag{B.9}$$

and in exponential form

$$e^{t\mathcal{A}} a(f)^* e^{-t\mathcal{A}} = a(e^{tA} f)^*. \tag{B.10}$$

Now, let $|0\rangle$ be the ground state of $\mathcal{F}^V$. For $n \in \mathbb{N}$, and $f_1, f_2, \ldots, f_n \in L^2(V)$, the finite linear combinations of the states

$$|f_1, f_2, \ldots, f_n\rangle = a(f_1)^* a(f_2)^* \cdots a(f_n)^* |0\rangle \tag{B.11}$$

are dense in $\mathcal{F}$, which is another way of stating that $|0\rangle$ is cyclic w.r.t. the algebra generated by the creation operators. Therefore, we need only test our assertion on vectors of the type (B.11). Using (B.10) with $t = 1$, and observing that $\mathcal{A}|0\rangle = 0$, we obtain

$$\begin{aligned} e^{\mathcal{A}}|f_1, \ldots, f_n\rangle &= e^{\mathcal{A}} a(f_1)^* e^{-\mathcal{A}} \cdots e^{\mathcal{A}} a(f_n)^* e^{-\mathcal{A}} |0\rangle \\ &= a(e^A f_1)^* \cdots a(e^A f_n)^* |0\rangle \qquad\qquad (B.12) \\ &= |e^A f_1, \ldots, e^A f_n\rangle. \end{aligned}$$

The existence of C is a consequence of the spectral theorem. We call $\mathcal{C}$ the corresponding quadratic form in a_i^*, a_j. Through the repeated use of (B.12), one checks that applying $e^{\mathcal{A}} e^{\mathcal{B}} e^{\mathcal{A}}$ to the states (B.11) is the same as applying $e^{\mathcal{C}}$. The semiboundedness of A and B ensures that the domain of their exponentials is the whole $L^2(V)$ and all quantities are well defined. Q.E.D.

B.3 Proof of Lemma 8.6

For any symmetric operator A, $\chi_\Lambda A \chi_\Lambda \leq \chi_\Lambda |A| \chi_\Lambda$. Hence $|\chi_\Lambda A \chi_\Lambda| \leq \chi_\Lambda |A| \chi_\Lambda$ and $\mathrm{Tr}|\chi_\Lambda A \chi_\Lambda| \leq \mathrm{Tr}_\Lambda |A|$. When $A = D_V$, the convergence of the trace is proven by writing the further estimate $\mathrm{Tr}_\Lambda |D_V| \leq \mathrm{Tr}_V |D_V|$ and then summing an integrable sequence of discrete eigenvalues. For $A = D$, one uses the Dirac-delta representation of the trace to find out that $\mathrm{Tr}_\Lambda |D| = |\Lambda| \, (2\pi)^{-d} \int |\widehat{d}|$. The first assertion of the lemma has been proven.

As for the second part, let us write

$$\begin{aligned} & \mathrm{Tr}(|\chi_\Lambda (D_V - D) \chi_\Lambda|) = \mathrm{Tr}(U \chi_\Lambda (D_V - D) \chi_\Lambda) \qquad (B.13) \\ = \;& \mathrm{Tr}_V (U \chi_\Lambda D_V \chi_\Lambda) - \mathrm{Tr}(U \chi_\Lambda D \chi_\Lambda) =: T_\ell - T, \end{aligned}$$

where U is the partial isometry $L^2(\Lambda) \longrightarrow L^2(\Lambda)$ that realizes the spectral decomposition as in [RS, Thm. IV.10]. It is convenient to use the position representation for the bases. So, $\psi^{(k)}(x) = e^{ik\cdot x}$ and, as defined in Chapter 7, $\psi_V^{(k)} = \psi^{(k)}\chi_V$. Let us work on T_ℓ: using the cyclicity of the trace one obtains

$$\begin{aligned} T_\ell &= \frac{1}{\ell^d} \sum_{k\in V'} \langle \psi_V^{(k)} | \chi_\Lambda U \chi_\Lambda D_V | \psi_V^{(k)} \rangle \\ &= \frac{1}{\ell^d} \sum_{k\in V'} \widehat{d}(k)\, \langle \psi_V^{(k)} | \chi_\Lambda U \chi_\Lambda | \psi_V^{(k)} \rangle \\ &= \frac{1}{\ell^d} \sum_{k\in V'} \widehat{d}(k)\, \langle \psi^{(k)} | \chi_\Lambda U \chi_\Lambda | \psi^{(k)} \rangle, \end{aligned} \tag{B.14}$$

the last equality being due to the presence of the indicator functions χ_Λ. In complete analogy with the above,

$$T = \frac{1}{(2\pi)^d} \int dk\, \widehat{d}(k)\, \langle \psi^{(k)} | \chi_\Lambda U \chi_\Lambda | \psi^{(k)} \rangle. \tag{B.15}$$

Since $|\langle \psi^{(k)} | \chi_\Lambda U \chi_\Lambda | \psi^{(k)} \rangle| \le |\Lambda|$, it is obvious that (B.14) tends to (B.15) for $\ell \to \infty$. Q.E.D.

B.4 Proof of Lemma 8.7

With regard to (8.23) and (8.26), $\det(1 + \widetilde{\zeta}\chi_\Lambda D_\varepsilon \chi_\Lambda)$ is entire in $\widetilde{\zeta}$ (hence in ζ) by [RS, Sec. XIII.17, Lemma 4(c)]. In order to evaluate its log (on the suitable Riemann surface) we need to avoid the zeros. Using [RS, Thm. XIII.106], we want to make sure that $\sigma(-\widetilde{\zeta}\chi_\Lambda D_\varepsilon \chi_\Lambda)$ does not hit 1. Step 3 in Section 8.2 (see in particular formulas (8.49)-(8.50) and the last paragraphs) shows that this is never the case if $\widetilde{\zeta} \notin (-\infty, -z^{-1} - 1]$, for FD, or $\widetilde{\zeta} \notin [z^{-1} - 1, +\infty)$, for BE. Q.E.D.

Actually, we can say more. Consider FD, just to fix the ideas. We see from (8.49) that the “floating” endpoint of $\sigma(-\widetilde{\zeta}\chi_\Lambda D_- \chi_\Lambda)$ is strictly contained in the segment $(0, -\widetilde{\zeta}/(1 + z^{-1}))$, which means that $\widetilde{\zeta}$ is allowed to exceed slightly $G_- - 1$, as given by (8.31), without any vanishing of (8.23). For bosons this is related to Lemma 8.1. In this case the “forbidden region” is $[\widetilde{\zeta}_{max}, +\infty)$, where $\widetilde{\zeta}_{max} := e^{\beta\lambda_{max}} - 1$ (see also the proof of Lemma 8.3).

In conclusion, for each finite Λ, the domain of analyticity of $\phi_{\varepsilon}^{\Lambda}(\zeta)$ is indeed strictly bigger than G_{ε}.

References

Part I: Infinite step billiards

[AA] V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, W. A. Benjamin, New York, 1968.

[AG] S. H. Aranson and V. Z. Grines, *On some invariants of dynamical systems on two-dimensional manifolds (necessary and sufficient conditions for the topological equivalence of transitive dynamical systems)*, Math. USSR Sbornik **19** (1973), 365-393.

[Ar] P. Arnoux, *Ergodicité generique des billards polygonaux*, Seminaire Bourbaki **696** (1988), 203-221.

[B] M. Boshernitzan, *Billiards and rational periodic directions in polygons*, Amer. Math. Monthly **99**(6) (1992), 522-529.

[BKM] C. Boldrighini, M. Keane and F. Marchetti, *Billiards in polygons*, Ann. Prob. **6** (1978), 532-540.

[CFS] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic theory*, Springer-Verlag, Berlin, 1982.

[DDL1] M. Degli Esposti, G. Del Magno and M. Lenci, *An infinite step billiard*, Nonlinearity **11**(4) (1998), 991-1013.

[DDL2] M. Degli Esposti, G. Del Magno and M. Lenci, *Escape orbits and ergodicity in infinite step billiards*, to appear in Nonlinearity.

[EFV] B. Eckhardt, J. Ford and F. Vivaldi, *Analytically solvable dynamical systems which are not integrable.*, Phys. **D 13**(3) (1984), 339-356.

[FK] R. H. Fox and R. B. Kershner, *Concerning the transitive properties of geodesics on a rational polyhedron*, Duke Math. J. **2** (1936), 147-150.

[GKT] G. Galperin, T. Krüger and S. Troubetzkoy, *Local instability of orbits in polygonal and polyhedral billiards*, Comm. Math. Phys. **169** (1995), 463-473.

[GiU] M.-J. Giannoni and D. Ullmo, *Coding chaotic billiards. I. Noncompact billiards on a negative curvature manifold*, Phys. **D 41**(3) (1990), 371-390.

[G1] E. Gutkin, *Billiard flows on almost integrable polyhedral surfaces*, Ergod. Th. & Dyn. Sys. **4** (1984), 569-584.

[G2] E. Gutkin, *Billiards in polygons*, Physica **D 19** (1986), 311-333.

[G3] E. Gutkin, *Billiards in polygons: survey of recent results*, J. Stat. Phys. **83** (1996), 7-26.

[GK] E. Gutkin and A. Katok, *Weakly mixing billiards*, in: Holomorphic Dynamics, X. Gomez-Mont, J. Seade, A. Verjovski, eds. LNM **1345**, Springer-Verlag, Berlin, 1989.

[Gu] M. C. Gutzwiller, *Chaos in classical and quantum mechanics*, Springer-Verlag, New York, 1990.

[H] N. E. Hurt, *Quantum chaos and mesoscopic systems*, Kluwer Academic, 1997.

[Ka] A. Katok, *Interval exchange transformations and some special flows are not mixing*, Israel J. Math. **35** (1980), 301-310.

[K] M. Keane, *Interval exchange transformations*, Math. Z. **141** (1975), 25-31.

[Ki] J. L. King, *Billiards inside a cusp*, Math. Intell. **17**(1) (1995), 8-16.

[KMS] S. Kerckhoff, H. Masur and J. Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. Math. **124** (1986), 293-311.

[Le] M. Lenci, *Escape orbits for non-compact flat billiards*, Chaos **6**(3) (1996), 428-431.

[L] A. M. Leontovich, *The existence of unbounded oscillating trajectories in a problem of billiards*, Sov. Acad. Sci., doklady Math. **3**(4) (1962), 1049-1052.

[Ma] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer-Verlag, Berlin, New York, 1986.

[M] H. Masur, *Closed trajectories of a quadratic differential*, Duke J. Math. **53** (1986), 307-314.

[PP] Ya. Pesin and B. Pitskel, *Topological pressure and the variational principle for noncompact sets*, Func. Anal. Appl. **18** (1984), 307-318.

[R] V. A. Rokhlin, *Lectures on the entropy theory of measure-preserving transformations*, Russ. Math. Surv. **22**(5) (1967), 1-52.

[S] Ya. G. Sinai (ed.), *Dynamical systems II (Ergodic theory with applications to dynamical systems and statistical mechanics)*, Encyclopaedia of Mathematical Sciences, 2. Springer-Verlag, Berlin, 1989.

[T] S. Tabachnikov, *Billiards*, Société mathématique de France, Paris, 1995.

[Tr] S. Troubetzkoy, *Billiards in infinite polygons*, Nonlinearity **12** (1999), 293-311.

[Tr2] S. Troubetzkoy, *Topological transitivity of infinite step billiards*, preprint (1999).

[V1] W. A. Veech, *Interval exchange transformations*, J. Analyse Math. **33** (1978) 222-272.

[V2] W. A. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. Math. **115** (1982) 201-242.

[W] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, GTM **79**, New York, 1982.

[ZK] A. Zemlyakov and A. Katok, *Topological transitivity of billiards in polygons*, Math Notes **18** (1975), 760-764.

Part II: Large deviations for quantum ideal systems

[ABV] N. Angelescu, J. G. Brankov and A. Verbeure, *General One-Particle Fluctuations of the Ideal Bose Gas*, J. Phys. A **29** (1996), 3341-3356.

[B] R. Balian, *From Microphysics to Macrophysics*, Vols. I and II, Springer-Verlag, Berlin-Heidelberg, 1992.

[BR] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vols. I and II, 2nd edition, Springer-Verlag, New York, 1997.

[BMV] M. Broidioi, B. Momont and A. Verbeure, *Lie Algebra of Anomalously Scaled Fluctuations*, J. Math. Phys. **36**(12) (1995), 6746-6757.

[C] J. T. Cannon, *Infinite Volume of the Canonical Free Bose Gas States on the Weyl Algebra*, Comm. Math. Phys. **29** (1973), 89-104.

[DLP] T. C. Dorlas, J. T. Lewis and J. V. Pulé, *The Full Diagonal Model of a Bose Gas*, Comm. Math. Phys. **156**(1) (1993), 37-65.

[E] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York, 1985.

[F] M. E. Fisher, *The Free Energy of a Macroscopic System*, Arch. Rational Mech. Anal. **17** (1964), 377-410.

[FL] M. E. Fisher and J. L. Lebowitz, *Asymptotic Free Energy of a System with Periodic Boundary Conditions*, Comm. Math. Phys. **19** (1970), 251-272.

[G] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, W. de Gruyter, Berlin-New York, 1988.

[Gi] J. Ginibre, *Some Applications of Functional Integration in Statistical Mechanics*, in: Statistical Mechanics and Quantum Field Theory, C. DeWitt, R. Stora (eds.), Gordon and Breach, New York, 1971.

[GVV] D. Goderis, A. Verbeure and P. Vets, *Dynamics of Fluctuations for Quantum Lattice Systems*, Comm. Math. Phys. **128** (1990), 533-549.

[H] K. Huang, *Statistical Mechanics*, J. Wiley & Sons, New York, 1987.

[LLS] J. L. Lebowitz, M. Lenci and H. Spohn, *Large Deviations for Ideal Quantum Systems*, preprint (1999).

[LZP] J. T. Lewis, V. A. Zagrebnov and J. V. Pulé, *The Large Deviation Principle for the Kac Distribution*, Helv. Phys. Acta **61** (1986), 1063-1078.

[O] S. Olla, *Large Deviations for Gibbs Random Fields*, Probab. Theory Related Fields **77**(3) (1988), 343-357.

[RS] M. Reed, B. Simon, *Methods of Modern Mathematical Physics*, Vols. I–IV, Academic Press, San Diego, 1972-1978.

[Ro] D. W. Robinson, *The Thermodynamic Pressure in Quantum Statistical Mechanics*, LNP 9, Springer-Verlag, New York, 1971.

[R] D. Ruelle, *Statistical Mechanics: Rigorous Results*, W. A. Benjamin, New York, 1969.

[SW] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, NJ, 1971.

[T] C. J. Thompson, *Mathematical Statistical Mechanics*, Macmillan, New York, 1971.

[Ti] E. C. Titchmarsh, *The Theory of Functions*, 2nd edition reprinted, Oxford University Press, Oxford-New York, 1968.

[vLL] M. van den Berg, J. T. Lewis and M. Lunn, *On the General Theory of Bose-Einstein Condensation and the State of the Free Boson Gas*, Helv. Phys. Acta **59** (1986), 1289-1310.

[vLP1] M. van den Berg, J. T. Lewis and J. V. Pulé, *A General Theory of Bose-Einstein Condensation*, Helv. Phys. Acta **59** (1986), 1271-1288.

[vLP2] M. van den Berg, J. T. Lewis and J. V. Pulé, *Large Deviations and the Boson Gas*, in: Stochastic Mechanics and Stochastic Processes, A. Truman, I. M. Davies (eds.), LNM 1325, Springer-Verlag, Berlin-Heidelberg, 1988.

[vFS] A. C. D. van Enter, R. Fernández and A. D. Sokal, *Regularity Properties and Pathologies of Position-Space Renormalization-Group Transformations: Scope and Limitations of Gibbsian Theory*, J. Stat. Phys. **72**(5-6) (1993), 879-1167.

[Wa] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1922.

[W] J. Weidmann, *Linear Operators in Hilbert Spaces*, Springer-Verlag, New York, 1980.

[ZUK] R. M. Ziff, G. E. Uhlenbeck and M. Kac, *The Ideal Bose-Einstein Gas, Revisited*, Phys. Rep. **32C** (1977), 169-248.

Vita

Marco Lenci

1988-93 Undergraduate student in Physics, Università di Bologna, Italy.

1993 Laurea in Physics, Summa Cum Laude, Università di Bologna.

1993-97 Graduate student in Mathematics, Università di Bologna.

1995-97 Visiting Fellow, Department of Mathematics, Princeton University.

1996 *Escape orbits for non-compact flat billiards*, Chaos **6**, 428-431.

1996 *Ergodic properties of the quantum ideal gas in the Maxwell-Boltzmann statistics*, Jour. Math. Phys. **37**, 5136-5157.

1996-99 Graduate student in Mathematics, Rutgers University.

1996-99 Teaching Assistant, Department of Mathematics, Rutgers University.

1998 *An infinite step billiard*, with M. Degli Esposti and G. Del Magno, Nonlinearity **11**, 991-1013.

1998 Ph.D. in Mathematics, Università di Bologna.

1999 *Escape orbits and ergodicity in infinite step billiards*, with M. Degli Esposti and G. Del Magno, to appear in Nonlinearity.

1999 *Large deviations for ideal quantum systems*, with J. L. Lebowitz and H. Spohn, preprint.

1999 Ph.D. in Mathematics, Rutgers University.