de Gruyter Textbook

Arno Berger

Chaos and Chance

An Introduction to Stochastic Aspects of Dynamics

Arno Berger

Chaos and Chance

An Introduction to Stochastic Aspects of Dynamics

Walter de Gruyter
Berlin · New York 2001

Author
Arno Berger
Institute of Mechanics
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Vienna
Austria

Mathematics Subject Classification 2000: 37-01; 37A50, 37A30, 70K55
Keywords: Bifurcation, Attractor, Chaos, Ergodic theorems, Markov operator, Markov chain

∞ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

Library of Congress — Cataloging-in-Publication Data

Berger, Arno, 1968–
Chaos and chance : an introduction to stochastic aspects of dynamics / Arno Berger.
p. cm.
Includes bibliographical references and index.
ISBN 3-11-016990-8 (pbk.). – ISBN 3-11-016991-6 (hardcover)
1. Differentiable dynamical systems. 2. Chaotic behavior in systems. I. Title.
QA614.8 .B48 2001
514'.352–dc21 2001042161

Die Deutsche Bibliothek — Cataloging-in-Publication Data

Berger, Arno:
Chaos and Chance : an introduction to stochastic aspects of dynamics / Arno Berger. – Berlin ; New York : de Gruyter, 2001
(De Gruyter textbook)
ISBN 3-11-016990-8 brosch.
ISBN 3-11-016991-6 Gb.

Printed in Germany.
Cover design: Hansbernd Lindemann, Berlin.
Typesetting using the author's TEX files: I. Zimmermann, Freiburg.
Printing and Binding: WB-Druck GmbH & Co., Rieden/Allgäu.

de Gruyter Textbook

Berger · Chaos and Chance

Preface

Since the nineteen-seventies *chaos* has become a highly popular term which, possibly due to its inherent vagueness, provides a succinct notion for complexity and unpredictability of dynamical systems. Consequently, there is no lack of textbooks that are concerned with the stunning aspects of *chaotic dynamical systems*, and implications thereof in science, engineering and economics. On the other hand, the statistically oriented approach to dynamics has also flourished during the past few decades, as for example can be seen from a number of excellent textbooks on *ergodic theory* and its applications. Dating back to the first half of the last century, this approach in some sense introduces the notion of *chance* even to completely deterministic systems. By bringing the two viewpoints together, an illustrative impression of the both natural and fruitful interplay of chaos and chance, or, more formally, of the geometrical and the statistical approach to dynamics may be gained. So far, however, a pertinent introductory though mathematically reliable text has been hard to find. It is the aim of the book in hand to help close this notable gap.

In writing this book it has been my intention to draw the reader's attention to several interesting examples, thereby motivating more general (and abstract) notions as well as results in a setting as simple as possible. Illuminating dynamical systems abound, and it need not be difficult to get an intuitive feeling of the complexity inherent in them. As one class of examples among others we shall repeatedly and on different levels of sophistication deal with billiards. By its conceptual simplicity nothing could be more deterministic and hence predictable than a billiard, could it? Given a specific shape of a table we shall, however, observe that the future fate of a voyaging billiard ball may be completely unpredictable beyond a surprisingly small number of reflections. In this case, calculation-based predictions of the ball's further journey are doomed to be no more reliable than throwing a die. But how does *chance* emerge from a purely deterministic system, and how may it help us to better understand the latter? It will take us some time to conceptualize and thoroughly address these questions.

It is my firm belief that illustrative examples are indispensable for developing a general theory, not only as initial justification but also as lasting motivation. Accordingly, it has never been my intention (let alone ability) to present a comprehensive monograph on the topic. As mentioned earlier, a host of excellent advanced texts can be drawn on for more extensive and detailed study; Appendix B contains a short and somewhat biased list of books to this purpose. I believe, however, that the elementary insights gained here together with the mathematical tools developed will enable the reader to study these advanced texts more rewardingly and with greater relish.

This book comprises five chapters and two appendices containing background material and references to the literature. As Chapter One serves as an informal introduction and Chapter Five lets the reader take a look at more advanced topics, Chapters Two to Four should be considered the core part of this text. Chapter Two briefly introduces the topologically oriented approach to chaotic dynamics. Chapters Three and Four focus on the statistical description of dynamical systems in some more detail. Especially, their resemblances to and differences from special stochastic, i.e. chance-driven and hence explicitly random, processes are thoroughly discussed. As throughout, the emphasis is on specific examples rather than on general results. By and by the reader will thus come to think of chaos and chance as of two sides of the same coin.

This is primarily a mathematics text. In the first place it addresses advanced undergraduates and beginning graduates with a sound knowledge of calculus. Ideally, the reader should also be endowed with a basic knowledge of measure theory. Appendix A gathers the most relevant notions and results from measure theory that we shall rely on in the main text; anyone to whom the content of Appendix A seems familiar is certainly well prepared. Another desirable prerequisite concerns discrete dynamical systems: though essentially self-contained, Chapter Two proceeds at a good pace, which may be demanding for a complete novice. To better appreciate its content, a precedent exposition to an introductory book, e.g. parts of Devaney's highly readable text, may be helpful. Yet a lack of either of these desirable prerequisites should not completely discourage the aspirant. Especially students from applied sciences are often highly motivated to learn more about the mathematics behind the systems they encounter in their respective discipline. In fact, I feel confident that this text will be useful to these readers too, provided they either take for granted the presupposed mathematical facts or (even better) look them up in a textbook and thus enhance their knowledge. With a purpose in mind, investing in one's mathematical skills is certainly worthwhile at any level of proficiency!

In order to keep the text fluent, calculations and considerations which I consider elementary are often stated in a brief form or skipped altogether. On a first or cursory reading one may well proceed by just taking note of the facts presented, without pondering on each of them. The conscientious reader, however, will use paper and pencil in order to work out and carefully check the steps condensed into phrases like *a straightforward calculation confirms*, etc. This is particularly relevant to comprehending the proofs: although these are really complete proofs containing all the essential ideas, it may often be advisable to thoroughly reconsider the details in order to get a fully elaborated argument. Exercises have been added in moderate number to each chapter: while some of them are routine, others are challenges for the curious, providing possible points from where to launch further studies. Throughout the whole book, the reader may thus decide to what extent to delve into the matter. For all I know, it is only personal commitment and interest covering also the sometimes technical details that leads to a truly working knowledge of the field. Accordingly, I hope that the reader will gain some inspiration from thoroughly pondering on specific problems of dynamics. May you feel inspired to consult more advanced references and to penetrate more and more the fascinating and incredibly multifarious world of dynamical systems!

This book has grown out of courses I have repeatedly given at Vienna's University of Technology. In turning the original notes into a serious textbook I received help and advice from quite a number of people. To all of them I feel sincerely grateful. In

particular, I was lucky to learn a lot about dynamics from Martin Blümlinger, Klaus Schmidt and Peter Szmolyan, who have let me benefit from their respective deep knowledge of the field. Peter also read parts of the manuscript and provided a host of corrections and improvements. For advice of both stylistic and general nature I am enormously indebted to Wuddy Grienauer as well as to Paul G. Seitz. The splendid illustration of *Snakes and Ladders* was kindly made available to me by Josie Porter and Tim Walters. Last but not least, I feel deep gratitude to Manfred Karbe for the most friendly and encouraging attention he patiently gave to this project.

Vienna, August 2001 *Arno Berger*

Contents

Chapter 1
Introduction

Even though defined in purely deterministic terms, dynamical systems from physics, biology, economics and other applied sciences may behave in quite a complicated and unpredictable manner, so that when looking at such systems we could easily get the impression of watching an experiment of *chance*. However, as there is no notion of chance in a completely deterministic setting, a seemingly erratic behaviour leads to several fundamental questions. How does uncertainty come about in this context and how does it affect our understanding of a system as a whole? What is the precise meaning of the word *chance* and, most important, how can we describe its emergence and implications? To pursue these questions and to find at least partial answers is the aim of this book. On our way we shall encounter numerous examples which will allow us to grasp the relevance of general notions and results. As this book primarily is about mathematics, several technicalities will have to be elaborated in some detail later in the text. In this introductory chapter, however, we are going to informally discuss a number of simple examples that will serve as starting-points and motivation for our further studies. The main point here is to observe how statistical and probabilistic aspects become important for the analysis of dynamical systems.

1.1 Long-time behaviour of mechanical systems

Very simple mechanical systems may evolve to a surprisingly complex behaviour if only they are observed for a sufficiently long period of time. We shall see and analyse such types of dynamics in a host of examples throughout this book. To get a first impression of the phenomena that may occur, let us start by looking at several *billiard systems*. A couple of fundamental notions may comfortably be introduced in the realm of these systems.

To explain what we mean by a billiard system let us consider a bounded open set Ω (the "table") in the plane which is connected and has a piecewise smooth boundary $\partial\Omega$. If we now put a small billiard ball somewhere on the table, make it move off with velocity $v \in \mathbb{R}^2$, $v \neq 0$ and assume that there is no loss of mechanical energy whatsoever (no friction, perfectly elastic reflection at the boundary etc.), then the resulting motion of the ball is easily described: inside Ω it moves with constant speed, whereas at the boundary it is reflected in such a way that *the angle of incidence equals the angle of reflection*. Since there is no loss of kinetic energy this motion goes on forever.

Let us first assume that the table Ω is rectangular with side-lengths a and b, respectively (see Figure 1.1). If we make the ball move off at time $t = 0$ with an angle α according to Figure 1.1, can we predict its position $x(t)$ for any $t > 0$? In fact, nothing could be simpler than this: observing that the projection of the ball's motion onto the x_1- and x_2-axis is given by a one-dimensional straight movement between walls with velocities $\pm v \cos\alpha$ and $\pm v \sin\alpha$, respectively, we could explicitly write down a formula for $x(t)$. It goes without saying that the angle α has to be chosen in such a way that the ball never hits a corner; most choices of α will satisfy this condition. There are, however, more challenging questions to ask: Will the trajectory eventually close, thus giving rise to a periodic motion? What will happen, if the trajectory fails to close? Can we – in the latter case – find a region of the table which is never hit by the ball?

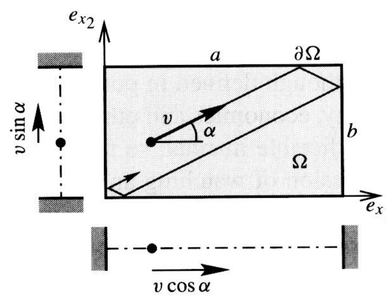

Figure 1.1. The rectangular table

To answer these questions we put our billiard system into a more tractable form by making the table twice as long and wide (cf. Figure 1.2). To benefit from this doubling procedure we consider straight motions with constant velocity on the doubled table subject to the following identification rule: for $0 \le x_1 \le 2a$ we consider the points $(x_1, 2b)$ and $(x_1, 0)$ as two guises of the same point; analogously we identify $(2a, x_2)$ and $(0, x_2)$ for all $0 \le x_2 \le 2b$. Concerning our straight motion this just says that we reenter from the bottom ($x_2 = 0$) if we have gone through the top ($x_2 = 2b$) and, analogously, reenter from the left if we have run out at the right. By means of

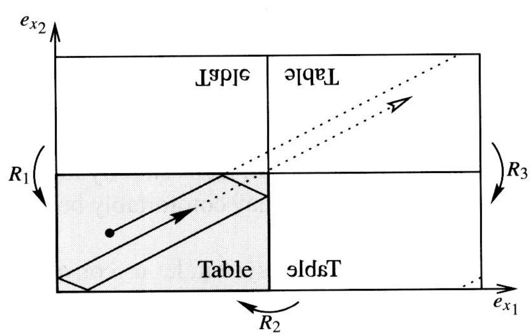

Figure 1.2. Each billiard trajectory uniquely corresponds to a straight line on the doubled table.

this identification and a repeated application of the reflections R_1, R_2, R_3 indicated in Figure 1.2 it is easy to see that each billiard trajectory within the rectangle Ω uniquely corresponds to a trajectory of the straight motion on the doubled table. Thereby we realize that the trajectory on the original table will close if and only if

$$p\,a \sin\alpha - q\,b\cos\alpha = 0 \quad \text{for some } (p,q) \in \mathbb{Z}^2,\ (p,q) \neq (0,0)\,.$$

Periodicity will therefore be observed if either $a/b \tan\alpha$ is a rational number or else $\cos\alpha = 0$. What happens if $a/b \tan\alpha$ is irrational? We cannot answer this question right now, but later we shall see that in this case every trajectory fills the rectangular table densely. In fact even more is true: we shall prove in Chapter Three that for $a/b \tan\alpha \in \mathbb{R}\backslash\mathbb{Q}$ the relation

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbf{1}_Q\big(x(t)\big)\,dt = \frac{\text{area}(Q)}{ab} \tag{1.1}$$

holds for any rectangle $Q \subseteq [0,a]\times[0,b]$; here $\mathbf{1}_Q$ denotes the *indicator function* of Q (see Appendix A), so that the integrand equals one precisely if $x(t) \in Q$ and zero otherwise. Observe that (1.1) essentially is a statistical statement about the behaviour of the trajectory $\{x(t) : t \geq 0\}$. Indeed, since the quantity at the left is nothing else but the asymptotic relative frequency of the ball being in Q, relation (1.1) may be rephrased as follows: asymptotically, for a long time T of observation, the relative frequency ("probability") of the billiard ball finding itself in Q is given by the portion of the whole table that is covered by Q. The larger Q is, the more often the ball will be found there. This last result even holds for sets more general than rectangles. Yet is it really obvious that in the long run all the regions depicted in Figure 1.3 are visited with the same relative frequency?

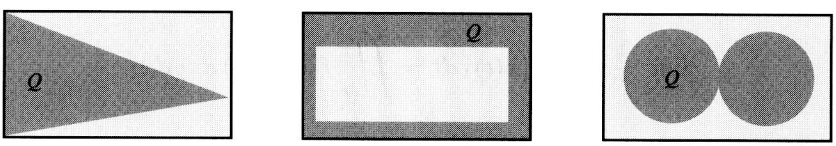

Figure 1.3. In the long run each shaded region Q is visited with the same relative frequency 0.5.

Let us now turn towards the *circular billiard*, that is the billiard on a disc, which probably has the simplest table with smooth boundary. From the left disc in Figure 1.4 it can be seen that the angle α enclosed by the billiard trajectory and the tangent to the boundary $\partial\Omega$ is the same at each point of reflection. As a consequence, the whole billiard trajectory is contained in the closed annulus

$$\overline{A}_\alpha := \{(x_1,x_2) : R^2\cos^2\alpha \leq x_1^2 + x_2^2 \leq R^2\}\,,$$

and each segment of the trajectory is tangent to the circle with radius $R|\cos\alpha|$ concentric to $\partial\Omega$ (see the billiard tables in Figure 1.4). As for the rectangular billiard there are two types of behaviour a billiard trajectory may exhibit: if $\alpha/\pi \in \mathbb{Q}$ then the trajectory will close after a finite number of reflections, thus yielding a periodic motion. If on

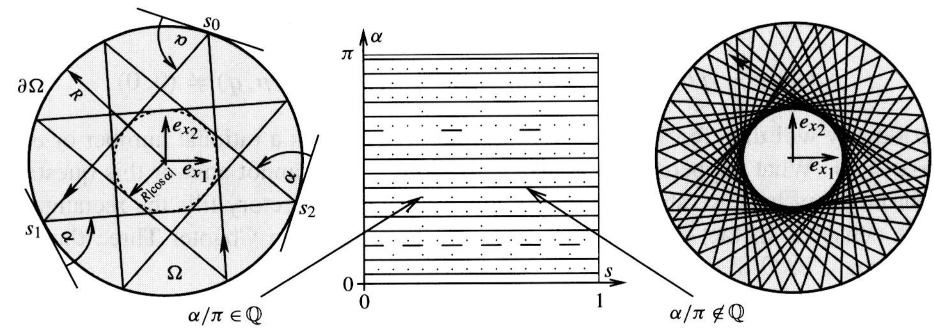

Figure 1.4. For the circular billiard the horizontal lines $[0, 1[\times\{\alpha\}$ are invariant under the associated billiard map T_{bill}.

the other hand α/π is irrational then the trajectory densely fills the annulus A_α; and we shall see that – in analogy to (1.1) –

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbf{1}_Q\big(x(t)\big)\,dt = \frac{1}{2\pi R\sin\alpha}\iint_Q \frac{dx_1\,dx_2}{\sqrt{x_1^2+x_2^2-R^2\cos^2\alpha}} \tag{1.2}$$

for every not-too-complicated set $Q \subseteq A_\alpha$ in this case.

Though elementary, a few observations are worth mentioning here. Firstly, (1.2) may be rewritten as

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbf{1}_Q\big(x(t)\big)\,dt = \iint_Q f_\alpha(x_1, x_2)\,dx_1\,dx_2$$

with the function $f_\alpha : A_\alpha \to \mathbb{R}$ defined as

$$f_\alpha(x_1, x_2) := \frac{1}{2\pi R\sin\alpha\sqrt{x_1^2+x_2^2-R^2\cos^2\alpha}}\,.$$

Evidently $f_\alpha \geq 0$ and $\int_{A_\alpha} f_\alpha(x_1, x_2)\,dx_1dx_2 = 1$; in probabilistic terms f_α therefore is a *density*. As we shall have occasion to observe again and again throughout this book, densities may be extremely useful for describing the long-term behaviour of dynamical systems.

A second observation is that a billiard system inside a table Ω whose boundary $\partial\Omega$ is a single smooth closed curve naturally induces a map on $\partial\Omega\times]0, \pi[$. To see this, parametrize $\partial\Omega$ by arc-length, take $(s, \alpha) \in \partial\Omega\times]0, \pi[$ and consider the billiard trajectory which emanates from the point with arc-length coordinate s and which encloses an angle α with the oriented local tangent (cf. Figure 1.4). It is natural then to assign to (s, α) the corresponding data of the next impact. Without loss of generality we may normalize $\partial\Omega$'s arc-length to one so that we end up with a *billiard map* T_{bill} which maps $[0, 1[\times]0, \pi[$ into itself. Since between any two reflection points the trajectory is just a straight line, the billiard map essentially contains all the information about the dynamics of the billiard system. In anticipation of our detailed studies in later chapters we now look at a few examples which indicate that the long-time behaviour of T_{bill} may differ considerably, depending on the specific shape of the table under consideration.

For the circular billiard it is straightforward to give an explicit formula for T_{bill}, namely

$$T_{\text{bill}} : (s, \alpha) \mapsto \left(s + \frac{\alpha}{\pi} (\text{mod } 1) \, , \, \alpha\right)$$

for all $(s, \alpha) \in [0, 1[\times]0, \pi[$. It is a remarkable feature of this map that it does not alter the second coordinate and thus maps the straight line $[0, 1[\times\{\alpha_0\}$ onto itself for any $\alpha_0 \in]0, \pi[$ (see Figure 1.4). The restriction $T_{\text{bill}}|_{[0,1[\times\{\alpha_0\}}$ may therefore be considered a map on $[0, 1[$; it is in fact a *rotation* and thus the simplest example of a *circle map*, a class of maps we shall deal with in Sections 2.4 and 3.2.

Admittedly, the circular billiard is easy to survey, and the need for a statistically oriented analysis thereof may not be too pressing. A slight modification of the table may, however, suffice to yield much more complicated dynamics of the associated billiard map. Consider for example the table depicted in Figure 1.5 which differs from a disk only by a rectangle inserted between the two halves of the disk. As can be seen from Figure 1.5 this innocent surgery yields a billiard system quite different from the circular one; traditionally it is referred to as the *stadium billiard* for its resemblance to the shape of an athletic field. As is indicated by Figure 1.5, typical trajectories of the stadium billiard do not show much regularity, neither do the orbits of individual points under the associated billiard map T_{bill}. In fact, a statistical approach is required to demonstrate that there *is* a certain regularity in the long-time behaviour of this specific billiard.

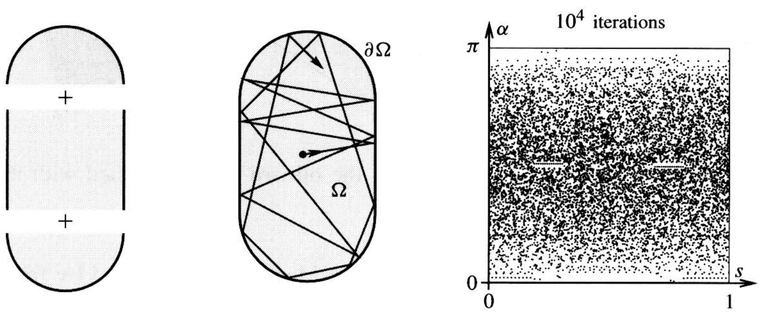

Figure 1.5. The *stadium billiard* and a typical orbit of T_{bill} (right)

The most basic statistical analysis certainly consists in drawing *histograms* of one or a few orbits of the billiard map T_{bill}. To this end, let us divide the space $[0, 1[\times]0, \pi[$ into a not-too-small number of squares S_i. Then we fix a point $x \in [0, 1[\times]0, \pi[$ and simply count how often $T_{\text{bill}}^k(x)$, i.e. the k-th iterate of x under T_{bill}, happens to fall into S_i for $0 \le k < n$. More formally, we numerically evaluate the relative frequencies

$$a_i(x,n) := \frac{\#\{0 \le k < n : T_{\text{bill}}^k(x) \in S_i\}}{n \ \text{area}(S_i)}$$

for various values of n. What help can the quantities a_i be for understanding the dynamics of T_{bill}? Most basically, $a_i \ge 0$ and

$$\sum_i a_i(x,n) \ \text{area}(S_i) \equiv 1 ,$$

so that we may interpret the family (a_i) as an approximation of a density. The larger a_i, the more often the iterates of x will visit S_i. Finally, we expect these quantities to tell us something about the long-time behaviour of T_{bill} for $n \to \infty$. How will $a_i(x,n)$ evolve as $n \to \infty$?

It turns out that the specific choice of x does not much affect the striking result displayed in Figure 1.6: in the long run each of the quantities a_i seems to converge! After we will have developed the statistical point of view in Chapter Three we shall not only be able to confirm this visual impression but also to explain the specific shape of the limit in Figure 1.6. Additionally, we shall see how these insights lead to a satisfactory description of the stochasticity inherent to the stadium billiard.

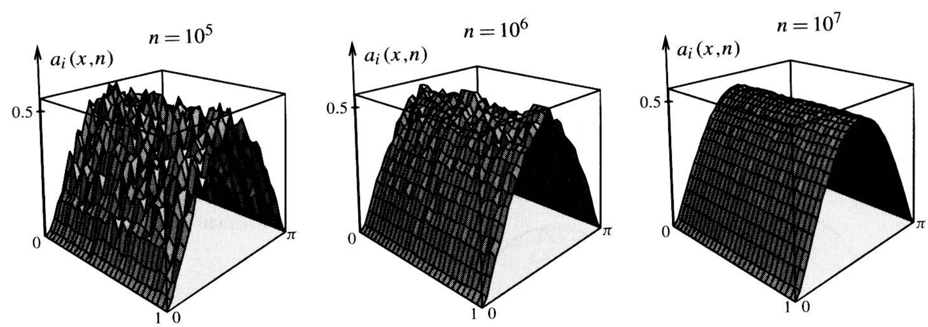

Figure 1.6. Empirical histograms drawn from the billiard map associated with the stadium billiard

As a final example of a billiard system consider a table bounded by four quarters of a circle curved *inward* (see Figure 1.7). Observe that this table is not convex. As a consequence, T_{bill} no longer is a continuous map, and it is not defined at the vertices of the table. Ignoring these technical difficulties for the time being we can perform a similar analysis as above, and we find that this billiard is by no means simpler than the stadium billiard (Figure 1.7). In fact, the mechanism which we can see at work here consists in an exponential instability of individual trajectories resulting in a *sensitive dependence on initial conditions* of the whole system. It is this latter effect which nowadays is prevalently considered as an essential ingredient of the notion of *chaos*, especially among physicists. Later we shall give a rigorous formalization, but for the specific billiard considered here we may informally discuss right away what sensitive dependence on initial conditions can mean for practical purposes.

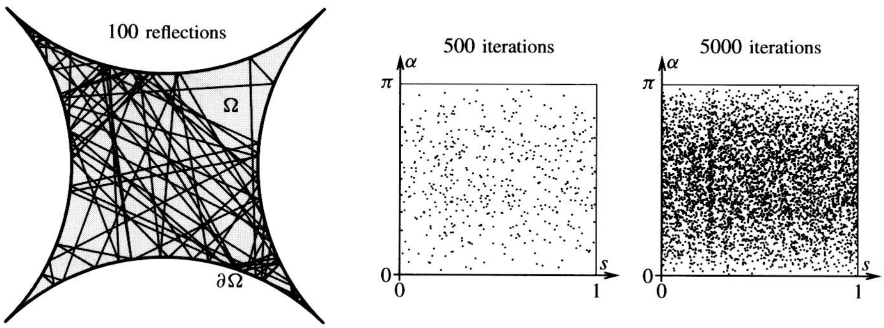

Figure 1.7. A trajectory for the dispersing billiard (left); the corresponding T_{bill}-orbit is likewise irregular.

Consider two billiard trajectories emanating from the same point in slightly different directions. Let δ denote the angle between these two directions; we consider δ as a perturbation parameter which is very small in absolute value. As a result of this small perturbation the two trajectories hit the boundary $\partial\Omega$ at different points which are $c\delta$ apart where c denotes a positive number depending on the local geometry of the boundary (see Figure 1.8). A heuristic argument shows that after reflection the directions of the trajectories will differ by an angle $(1 + 2c)\delta$. Since $|\delta|$ is very small this augmentation does not look critical. However, observe that the actual difference between the directions is multiplied by a factor larger than one at *each* reflection. As a dramatic consequence, small perturbations will grow more or less exponentially. In other words, even the smallest deviation δ will cause trajectories to significantly diverge after a frighteningly small number of reflections. Figure 1.8 provides a visualization of this effect with $\delta = 10^{-7}$: the two trajectories are close-by only for the first eleven reflections and then diverge completely.

Billiards like the one discussed here have been termed *dispersing*, a notion being self-explanatory by now. In the light of the above discussion we expect such billiards to behave chaotically in the long run. As we shall see in Chapter Three this is true in a precise sense, though working out the mathematical details is rather demanding.

Throughout this book we shall encounter and carefully analyse the effect of *sensitive dependence on initial conditions* for many systems. The practical impact of this effect has very clearly been seen already by the founders of modern dynamical systems theory.

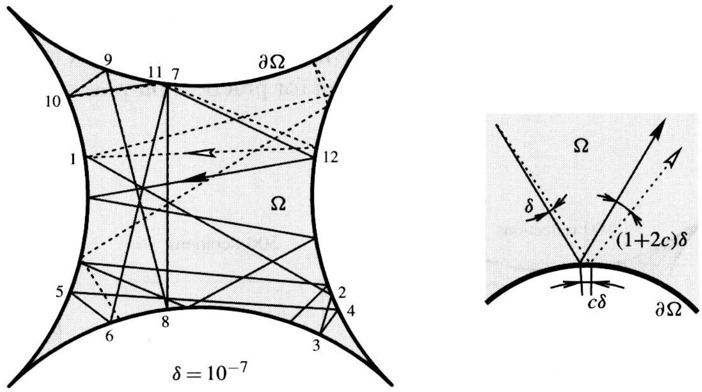

Figure 1.8. This billiard shows *sensitive dependence on initial conditions.*

In 1908 Poincaré wrote:

> *A very small cause, which escapes us, determines a considerable effect which we cannot ignore, and we then say that this effect is due to chance.*

As a consequence, our ability to predict the future evolution of individual trajectories (orbits) of such systems is tremendously limited. The best we can wish for in this situation is a meaningful *statistical* perspective of the dynamics, and we are going to develop this point of view from Chapter Three onward.

The last mechanical system we are going to introduce here has its roots in the work of Lagrange on celestial mechanics. Consider a system of $d+1$ points labelled $0, 1, \dots, d$ that perform a planar motion according to the following rule: the point 0 is fixed while for $k = 1, \dots, d$ the point labelled k circles around the point labelled $k-1$ with radius r_k and (absolute) angular velocity ω_k. One could think of a family of celestial bodies circulating around each other with constant angular velocities. What we are interested in here is the motion of the last point which may concisely be described by means of complex numbers according to

$$z(t) = a_1 e^{i\omega_1 t} + \ldots + a_d e^{i\omega_d t}$$

with $a_k \in \mathbb{C}$, $|a_k| = r_k$ for $k = 1, \dots, d$. As might be imagined this motion can be quite complicated and non-uniform (especially for large d; cf. Figure 1.9). Writing $z(t) = r(t)e^{i\varphi(t)}$, Lagrange asked whether an *(asymptotic) average angular velocity* $\omega_\infty := \lim_{t\to\infty} \varphi(t)/t$ could be assigned to that system. Here the angle $\varphi(t)$ is assumed to be continuous unless $z(t) = 0$; in the latter case it may exhibit a jump of an absolute value of at most π. If it exists at all, the quantity ω_∞ will describe on average the long-time behaviour of the system of rotating points. Lagrange found that