Symbolic dynamics: I. Finite dispersive billiards

Kai T Hansen†‡
Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

Received 14 February 1992, in final form 8 April 1993
Recommended by P Cvitanović

Abstract. Orbits in different dispersive billiard systems, for example the three-disk system, are mapped into a topological well-ordered symbol plane and it is showed that forbidden and allowed orbits are separated by a monotone pruning front. The pruning front can be approximated by a sequence of finite symbolic dynamics grammars.

PACS numbers: 0545

1. Introduction

The aim of this article is to find an effective description of all allowed orbits in a classical billiard system. The system considered is a point particle on a plane, moving freely and reflecting elastically off dispersing walls. In most of our examples the walls are the borders of circular shaped reflecting disks. Such systems have recently been the focus of much theoretical interest, because of their utility in studies of the quantum mechanics of classical chaotic systems [1–4]. Semiclassical calculations of the quantal spectrum of classically chaotic systems require summations over all periodic orbits of the classical system [5], and it is essential to have an effective description of the allowed classical orbits.

The solution we propose here is to apply techniques developed in the study of one-dimensional unimodal maps and two-dimensional Hénon-type maps. In the unimodal maps the ordering of allowed orbits along the parameter axis is given by the MSS theory [6], i.e. the kneading sequence of the critical point [7]. From the kneading sequence one constructs a binary number \( y \) and the interval \((y_0, 1)\) is the primary pruned region of the unimodal map. An orbit is forbidden if the orbit, or any of its images or pre-images, yields a binary number \( y \) in the pruned region \((y_0, 1)\). A convenient visualization of the allowed orbits is afforded by a binary tree on which every allowed orbit corresponds to a path down the tree. The forbidden orbits correspond to the forbidden branches, i.e. the branches that have to be pruned [8] from the tree of allowed orbits.

A unimodal map such as the logistic map has a parameter space window for each stable periodic orbit. Within such a window the kneading sequence is periodic, and one can find a finite number of forbidden symbol strings and construct a new unpruned alphabet [9]. One can also describe allowed strings as walks on a finite Markov graph constructed from the pruned binary tree [10]. Chaotic dynamics does not in general have a finite Markov partition, but we believe [11] that the statistical properties can be approximated from calculations on repellors corresponding to converging sequences of nearby windows with stable orbits.

† Permanent address: Department of Physics, University of Oslo, Box 1048, Blindern, N-0316 Oslo, Norway.
‡ e-mail: khansen@nbivax.nbi.dk
In [8] Cvitanović et al have extended these methods to two-dimensional maps like the Hénon map [12] and have shown that numerically one can obtain a connected primary pruned region in a two-dimensional space \((\gamma, \delta)\) constructed from the properly ordered symbolic future and the past of the orbits; the border of the area is the pruning front. The conjecture is that this curve is monotone, and that there are no other forbidden orbits than those that pass through the primary pruning region. One can approximate the grammar of forbidden orbits by forbidding the finite string that corresponds to rectangles in the symbol plane. Each such approximation gives a finite number of finite length forbidden symbol strings. From these strings one can make successively refined Markov partitions and obtain well-converging estimates of averages such as the topological entropy [13–15]. The pruning front is the symbol plane representation of the primary homoclinic tangencies introduced by Grassberger and Kantz [16] in order to partition the Hénon map non-wandering set into a ‘left’ and a ‘right’ side.

1.1. Symbols in billiard systems

In billiard systems the choice of symbols is much easier; a symbol for a bounce off each concave wall, for example a disk or a branch of the hyperbola, is natural and unambiguous [17–20]. The billiards considered here are hyperbolic systems with only unstable orbits; the stable and unstable manifolds of the orbits are never tangential to each other. An orbit in the billiard is unstable also at the point where it is pruned, in contrast to the Hénon map and smooth Hamiltonian potentials, where pruning always involves inverse bifurcations and stable orbits [21, 22].

The construction of billiard symbolic dynamics proceeds in two steps: (1) define a symbol plane that is topologically well ordered; and (2) determine the boundary orbits between the legal and illegal orbits. The well-ordered symbol plane is defined by introducing a new symbol alphabet and associating with every orbit a pair of coordinates constructed from the alphabet. A similar well-ordered symbolic alphabet has been introduced for a different scattering problem by Troll [23]. The second issue is simple. We show that for the billiards one starts a tangential bounce from every point on the wall and these are the boundary orbits.

The symbol plane we construct is complete for a system with a finite number of hard disks situated far from each other in the configuration space. A physically motivated example is a particle bouncing between four hyperbolas [4, 24]. This billiard does not have the obvious limit of a complete Cantor set but we use the same symbol plane as the four-disk billiard and orbits are pruned the same way.

2. Covering alphabet for dispersing billiards

For simplicity we choose to describe the covering alphabets for dispersing billiards by using billiards with a finite number of identical circularly shaped walls. In general the walls do not have to be circular, but they do have to be dispersing. When they are identical, the configuration space symmetry enables us to reduce the number of symbols and to simplify the problem.

If we take a billiard system with sufficiently separated disks in the plane there exists an orbit which visits the disks in any desired order.

Let the disks be enumerated by \(s \in \{1, 2, \ldots, N\}\) and denote the bounce of the point particle on disk \(s\) at the discrete time \(r\) by \(s_r\). Dispersing (concave) walls cannot have two
succeeding bounces off the same wall, i.e. the string \( ss_\ldots \) is forbidden, but all other strings are allowed for sufficiently separated disks.

Given the infinite symbolic past of the particle bouncing on disk \( s_0 \) at time \( t = 0 \), that is \( \ldots s_{-3} s_{-2} s_{-1} \), and the infinite symbolic future \( s_1 s_2 s_3 \ldots \), the position of the bounce \( x \) and the outgoing angle \( \phi \) measured from the normal vector of the wall are uniquely determined. The phase space coordinates we use are the position and the angle \((x, \phi)\) of the bounce off disk \( s_0 \).

If the \( N \)-disk system is open, then almost all points \((x, \phi)\) correspond to a particle that has entered the system from outside at some earlier time and is going to escape from the system at some later time. The union of all points contained within the billiard for all times is the non-wandering set \( \Lambda \). For a sufficiently separated disk billiard, \( \Lambda \) is a Cantor set in the phase space similar to the non-wandering set for the Smale horseshoe [25]. Each point in this Cantor set is described by a unique symbolic past and future. However, if we use the disk symbols \( s_i \) to describe a point in the Cantor set, there is no obvious ordering of the symbols from 'small' symbols to 'large' symbols that would reflect an increase of the position or angle coordinate in the phase space. A symbolic description with a natural ordering of this kind we call well-ordered symbols and we show how these can be constructed by a few examples. Other configurations of disks can be worked out in a similar way.

### 2.1. Three-disk symbolic dynamics

Let the billiard system consist of three equal disks with radius 1 and distance \( r \) between the centres, enumerated anticlockwise, as in figure 1. Without loss of generality we assume that the particle bounces off disk number 1 at time \( t = 0 \) (\( s_0 = 1 \)). Position \( x \) is the angle describing the bounce on disc 1 and \( \phi \) is the angle from the normal vector. Define \( \Lambda^t \) as the union of points in phase space that give at least \( t \) bounces in the billiard after bouncing off disk 1. Thus \( \Lambda^1 \) is all \((x, \phi)\) of disk 1 corresponding to a particle that bounces either off disk 2 or off disk 3 at time \( t = 1 \). \( \Lambda^1 \) consists of two diagonal strips in phase space, with \( \phi \) decreasing when \( x \) is increasing. The lower strip in the phase space is the union of all points in phase space corresponding to a particle bouncing off disk 2 at time \( t = 1 \) and the symbol string for this strip is \( s_0 s_1 = 12 \). The upper strip is the union of all points corresponding to a particle that bounces off disk 3 at time \( t = 1 \), and this strip has symbol string \( s_0 s_1 = 13 \). The order of the two strips with increasing \( x \) or \( \phi \) is \( s_0 s_1 = 12, 13 \). \( \Lambda^2 \) is four diagonal strips, two inside each of the two strips of \( \Lambda^1 \). The symbol description of the four strips ordered after increasing values of \( x \) or \( \phi \) is \( s_0 s_1 s_2 = 121, 123, 132, 131 \).

Table 1 shows the symbol strings for four bounces ordered after increasing values of \( x \) and \( \phi \). The limit of \( \Lambda^t \) when \( t \to \infty \) is a Cantor set of lines with a unique labelling \( s_0 s_1 s_2 \ldots \) for each line.

Let new symbols be defined from a combination of two symbols \( s_{i-1} \) and \( s_i \), and such that the ordering of the Cantor set lines in the phase space is preserved. In table 1 new symbols \( w_i \in \{0, 1\} \) are written together with the old symbols for the first four bounces. From the symbols \( w_1 w_2 w_3 \ldots \) a rational binary number \( \gamma = 0.w_1 w_2 w_3 \ldots = \sum_{i=1}^{\infty} w_i / 2^i \) is constructed. This symbolic coordinate \( \gamma \) increases with increasing values of \( x \) and increasing values of \( \phi \). To construct the symbols \( w_i \) let first \( v_i \) denote whether two consecutive bounces \( s_{i-1} \) and \( s_i \) take place in a clockwise or anticlockwise direction. The symbols \( v_i \) are not well ordered because a bounce in a dispersing (concave) wall reverses the ordering in the phase space. At each bounce one has to invert the symbols \( v_i \) and this gives the algorithm
Figure 1. The three disks in configuration space enumerated anticlockwise.

Table 1. Ordering of four bounces of a particle starting on disc 1 at time 0. The symbol \( s_t \) is the number of the disc, while \( w_t \) is a new symbol reflecting the ordering in phase space.

| \( s_0 \) | \( t=1 \) |
|---------|---------|
| 12121   | 0000    |
| 12123   | 0001    |
| 12132   | 0010    |
| 12131   | 0011    |
| 12313   | 0100    |
| 12312   | 0101    |
| 12321   | 0110    |
| 12323   | 0111    |
| 13232   | 1000    |
| 13231   | 1001    |
| 13213   | 1010    |
| 13212   | 1011    |
| 13121   | 1100    |
| 13123   | 1101    |
| 13132   | 1110    |
| 13131   | 1111    |

\[ v_t = s_t - s_{t-1} \]

if \( v_t < 1 \) then \( v_t = v_t + 3 \)

\[ w_t = \begin{cases} 
  v_t - 1 & \text{if } t \text{ odd} \\
  2 - v_t & \text{if } t \text{ even.}
\end{cases} \tag{1} \]

The new symbols are constructed to reflect the ordering of the lines \( \Lambda^+_\infty \) in the phase space. Exchanging odd and even in algorithm (1) gives symbols \( w'_t = 1 - w_t \), and a binary number \( \gamma' = 0.w'_1 w'_2 w'_3 \ldots \) decreasing with \( x \) and \( \theta \), and therefore also the symbols \( w'_t \) are well ordered. From an orbit described by the symbols \( s_t \) we have two mappings to the two well-ordered symbols \( w_t \) and \( w'_t \).

The particle bouncing at \( (x, \phi) \) also has a past, and a symbol string describing this past. Let \( \Lambda^-_t \) be the union of points in phase space corresponding to a particle that has at least \( t \) bounces in the billiard before arriving at \( (x, \phi) \). Then \( \Lambda^-_t \) is the union of all points arriving
at disk 1 after being bounced off disk 2 or off disk 3, and \( \Lambda^+_1 \) is two diagonal strips in the phase space. The incoming angle has opposite sign to the outgoing angle and the strips in \( \Lambda^+_1 \) have increasing \( \phi \) with increasing \( x \). We get the new symbols \( w_t \) from algorithm (1) with \( t < 0 \). The symbolic coordinate for the past is \( \delta = 0.w_0w_{-1}w_{-2} \ldots = \sum_{i=1}^{\infty} w_{-i}/2^i \). The coordinate \( \delta \) is increasing with increasing value of \( x \), but \( \delta \) is decreasing with increasing value of \( \phi \). If we choose the value \( \gamma' \) for the future we have \( \delta' = 0.w'_0w'_{-1}w'_{-2} \ldots \) for the past.

Let \( \Lambda_t = \Lambda^+_1 \cap \Lambda^-_1 \). Then \( \Lambda_2 \) consist of 16 = \((2 \times 2)^2\) areas in the phase space. Each of the sixteen disjoint sets of \( \Lambda_2 \) has a unique enumeration by the four binary symbols \{w_{-1}, w_0, w_1, w_2\}. The set \( \Lambda = \Lambda_\infty \) is a Cantor set and each point in the set is represented by a bi-infinite symbol string \( \ldots w_{-2}w_{-1}w_0w_1w_2 \ldots \). The set \( \Lambda \) is the non-wandering set of the billiard.

The symbolic plane is the unit square \((\gamma, \delta); 0 \leq \gamma, \delta \leq 1\). Each non-wandering orbit is represented by a point in this plane, and each point in the plane is one of the non-wandering orbits. As mentioned earlier, we actually get two points in the symbol plane from one orbit. Because of the symmetries of rotation and reflection in the billiard, we get a number of different orbits in the billiard from one point in the symbol plane, but all the orbits have the same length and stability. The symbolic plane is a representation of the phase space of the well-separated billiard, with all gaps in the Cantor set removed. This is a very convenient space to work in; the curved lines in the phase space become straight lines and this space does not change if the distance \( r \) changes or if the borders of the disks are slightly changed, as long as the symmetry and concavity are kept.

The billiard can be reduced to a fundamental domain [20]. The fundamental domain is one sixth of the original three-disk billiard and it is tiling the whole billiard. In our phase space the fundamental domain is the part with \( x \leq \frac{1}{3} \pi \) and is constructed in the symbol plane as follows. The symbol string \( \ldots s_{-2}s_{-1}s_0s_1 \ldots \) gives two symbolic coordinates \((\gamma, \delta)\) and \((\gamma', \delta')\), but if \( \gamma + \delta \neq 1 \) then one point is in the fundamental region and the other point is in the region \( x > \frac{1}{3} \pi \). We find the two points \((\gamma, \delta)\) and \((\gamma', \delta')\) and choose \((\gamma, \delta)\) if \( \gamma + \delta < 1 \) or choose \((\gamma', \delta')\) if \( \gamma + \delta > 1 \). If \( \gamma + \delta = 1 \) then both \((\gamma, \delta)\) and \((\gamma', \delta')\) are on the border of the fundamental domain, and as our convention we choose \((\gamma, \delta)\) if \( \gamma > \frac{1}{2} \) and choose \((\gamma', \delta')\) if \( \gamma < \frac{1}{2} \). If we have a billiard without symmetry we only use one map from \( s_t \) to \( w_t \) and we do not have any fundamental domain.

### 2.2. Symbolic dynamics for \( N \) disks on a circle

As a generalization of the three-disk billiard let \( N \) equal disks have the centre of each disk on a large circle and let the distance between centres of neighbour disks on the large circle be \( r \). Then \( \Lambda_1^+ \) is \((N - 1)\) strips in the phase space. The well-ordered symbols \( w_t \in \{0, 1, 2, \ldots, (N-2)\} \) are constructed from the anticlockwise enumeration of the disks \( s_t \in \{1, 2, \ldots, N\} \) using the algorithm

\[
\begin{align*}
v &= s_t - s_{t-1} \\
\text{if } v_t < 1 \text{ then } v_t &= v_t + N \\
w_t &= \begin{cases} 
  v_t - 1 & \text{if } t \text{ odd} \\
  N - v_t - 1 & \text{if } t \text{ even}
\end{cases} \quad (2)
\end{align*}
\]

and the opposite ordered symbols are \( w'_t = N - w_t - 2 \). When \( N = 3 \) this is the same algorithm as (1), and for \( N = 4 \) this is the algorithm in [24]. From \( w_t \) we
construct base \((N - 1)\) symbolic coordinates \(\gamma = \sum_{t=1}^{\infty} w_t/(N - 1)^t\) and \(\delta = \sum_{t=1}^{\infty} w_{t-1}/(N - 1)^t\) where \(0 \leq \gamma, \delta \leq 1\), and in a analogous manner we construct \(\gamma'\) and \(\delta'\).

The reduction to the fundamental domain in the symbol plane is the same as for three disks. If \(N\) is even there is one period 2 orbit with only one representation in the symbol plane \((\gamma, \delta) = (\gamma', \delta') = (\frac{1}{2}, \frac{1}{2})\).

### 2.3. \(N\) disks with a centre disk

Let the billiard be a configuration of \(N\) disks on a large circle as in the billiard above and in addition with one disk in the centre of this large circle. The radius of each disk is 1 and the distance between neighbour disks is \(r\). The disks on the circle are enumerated anticlockwise from 1 to \(N\) and a bounce off the disk in the centre is given the symbol \((N + 1)\). Figure 2 shows this configuration with \(N = 6\). If the number of disks on the large circle is even, then, because of the disk in the centre, a point particle cannot bounce between two disks opposite to each other on the large circle. From the \((N + 1)\) symbols of the disks, we get \((N - 1)\) well-ordered symbols, and \(A_{\text{w}}\) consists of \((N - 1)\) strips in the phase space. With \(N\) (even) disks on the large circle and disk number \((N + 1)\) in the centre of the large circle the algorithm defining the well-ordered symbols \(w_t \in \{0, 1, 2, \ldots, (N - 2)\}\) is

\[
\begin{align*}
\text{if } s_t &= (N + 1) \text{ then } w_t &= \frac{1}{2}(N - 2) \\
\text{else if } s_{t-1} &\neq (N + 1) \text{ then} \\
&\quad v_t = s_t - s_{t-1} \\
&\quad \text{if } v_t < 1 \text{ then } v_t = v_t + N \\
&\quad w_t = \begin{cases} 
&v_t - 1 & \text{if } t \text{ odd} \\
&N - v_t - 1 & \text{if } t \text{ even}
\end{cases} \\
\text{else if } s_{t-1} &= (N + 1) \text{ then} \\
&\quad v_t = s_t - s_{t-2} \\
&\quad \text{if } v_t < -\frac{1}{2}(N - 2) \text{ then } v_t = v_t + N \\
&\quad \text{if } v_t > \frac{1}{2}(N - 2) \text{ then } v_t = v_t - N \\
&\quad w_t = \begin{cases} 
&\frac{1}{2}(N - 2) + v_t & \text{if } t \text{ odd} \\
&\frac{1}{2}(N - 2) - v_t & \text{if } t \text{ even}.
\end{cases}
\end{align*}
\]

The configuration with \(N = 6\) can be looked at as a first step toward a description of the Lorentz gas [26, 27], a triangular lattice with a hard disk in each lattice point and a point particle scattering in the lattice.

If the number of disks \(N\) on the large circle is odd, a point particle can reach all other disks after bouncing off one disk when the disks are sufficiently separated. The algorithm giving the symbols \(w_t \in \{0, 1, 2, \ldots, (N - 1)\}\) is

\[
\begin{align*}
\text{if } s_t &= (N + 1) \text{ then } w_t &= \frac{1}{2}(N - 1) \\
\text{else if } s_{t-1} &\neq (N + 1) \text{ then} \\
&\quad v_t = s_t - s_{t-1}
\end{align*}
\]
Figure 2. The $6 + 1$ disks, enumerated anticlockwise, with $r = 2.2$. The orbit is tangent to disk 7, and therefore belongs to the pruning front for the system.

\[
\text{if } \nu_t < 1 \text{ then } \nu_t = \nu_t + N \\
\nu_t = \begin{cases} 
\nu_t - 1 & \text{if } t \text{ odd} \\
N - \nu_t - 1 & \text{if } t \text{ even}
\end{cases} \\
\text{if } \nu_t > \frac{1}{2}(N - 1) \text{ then } \nu_t = \nu_t + 1
\]

\text{else if } s_{t-1} = (N + 1) \text{ then}

\nu_t = s_t - s_{t-2}

\text{if } \nu_t < -\frac{1}{2}(N - 1) \text{ then } \nu_t = \nu_t + N

\text{if } \nu_t > \frac{1}{2}(N - 1) \text{ then } \nu_t = \nu_t - N

\nu_t = \begin{cases} 
\frac{1}{2}(N - 1) + \nu_t & \text{if } t \text{ odd} \\
\frac{1}{2}(N - 1) - \nu_t & \text{if } t \text{ even}
\end{cases}
\]

2.4. Orientation exchange at a bounce

One can explain why all the algorithms reverse the ordering of symbols at each bounce by a detailed description of a bounce. Assume that two particles move along two arbitrarily close parallel lines and bounce off the dispersing (concave) wall at an angle $\phi \neq 0$, as in figure 3. The particle that hits the wall first bounces off the wall along a new direction. The second particle crosses the first outgoing line and then bounces off the wall. The two outgoing lines cannot cross each other as long as the wall is concave or straight. The particle which is on the left-hand side before the bounce moves on to the right-hand side after the bounce, and this causes a change in the orientation at the bounce. If the angle $\phi = 0$, the orbits do not cross but, because the velocity is reversed, the orbit that was on the left-hand side before the bounce is on the right-hand side after the bounce.

For a bounce off a focusing (convex) wall, as in figure 4, the two parallel orbits cross twice if $\phi \neq 0$. There is one crossing after the first particle has bounced off the wall and one crossing after the second particle has bounced off the wall. The particle which is on the left-hand side before the bounce is on the left-hand side also after the bounce, and the
orientation does not change. This is true also for $\phi = 0$ when the orbits cross once after the bounces. These bounces take place, for example for a point particle moving in the stadium billiard. Orbits bouncing only off the semicircular parts of the stadium walls can therefore be described by an ordered rotation number [28]. The descriptions of all the orbits in the stadium billiard in well-ordered symbols are given in [29].

2.5. A condition for sufficiently separated disks

We want a condition on the geometrical construction of the non-wandering set to distinguish between sufficiently separated disks and the case of pruning. We find that the disks are sufficiently separated if condition 1 below is true.

The definition of $\Lambda^+_1$ in the general case is as follows: $\Lambda^+_1$ consists of $M$ strips and each strip $m^+$ is the union of points $(x, \phi)$ from which a straight line starting at $x$ with angle $\phi$ hits a particular disk. As a straight line may go through other disks the $M$ strips are not necessarily disjoint. In the same way, we define $\Lambda^-_1$ as consisting of $M$ strips where each strip $m^-$ is the union of points $(x, \phi)$ where a line from point $x$ with angle $-\phi$ hits a particular disk. We call the intersection of a strip $m^+$ and a strip $m^-$ a rectangle. The set $\Lambda_1 = \Lambda^+_1 \cap \Lambda^-_1$ then consists of $M^2$ rectangles not necessarily disjoint. The construction of $\Lambda_T$ with $T > 1$ follows from demanding that the outgoing angle at a bounce is equal to the incoming angle, but allowing the straight lines to go through a disc. Then $\Lambda_T$ consists of $M^{2T}$ rectangles.

Condition 1. There exists a number $0 < T < \infty$ such that $\Lambda_T$ consists of $M^2$ disjoint areas where each area is inside one of the $M^2$ rectangles of $\Lambda_1$.

The iteration of the $M^2$ disjoint areas corresponds to one more bounce and gives the result that each of the $M^2$ disjoint areas contains $M^2$ new disjoint areas. The $M^4$ rectangles of $\Lambda_2$ then contain $M^4$ disjoint areas. By induction we find that $\Lambda_{T+T'}$ gives $M^{2T}$ disjoint areas inside the $M^{2T}$ rectangles of $\Lambda_T$. From this it follows that, even if the rectangles of $\Lambda_T$ overlap each other, the parts of the non-wandering set belonging to the rectangles do not overlap. A symbol string $(w_{-T+1}w_{-T+2} \ldots w_{T-1}w'_T)$, with $w_j \in \{0, 1, \ldots, (M-1)\}$, uniquely corresponds to one rectangle when describing the part of the non-wandering set in this rectangle. It then follows that the disks are sufficiently separated.
If there does not exist a $T$ according to the condition it may be that an infinitesimal change of the parameter gives a $0 < T < \infty$ and this is the critical parameter value where pruning starts. If none of the above is true, then the rectangles always overlap and there is pruning in the system and not all symbol strings correspond to an orbit.

3. Pruning

When the distance $r$ between the disks is small, a subset of the orbits is forbidden (not admissible). We conjecture that only two kinds of forbidden orbit exist in the dispersing billiards without corners: orbits passing through a forbidden region (e.g. through a disk) and orbits going into a forbidden region and bouncing off the wall from the focusing side. Figure 5(a) shows a part of two legal orbits, one passing the dispersing wall and one bouncing off the wall on the dispersing side. Figure 5(c) shows part of two forbidden orbits, one passing through the dispersing wall and the other going into the forbidden region and bouncing off the wall from the focusing side. Figure 5(b) shows that the limiting orbit of both orbits is a line tangential to the disk. An orbit is therefore pruned together with at least one other orbit.

![Figure 5](image)

Figure 5. Allowed, boundary and forbidden orbits. (a) Two legal orbits, (b) two orbits on the pruning front, (c) two forbidden orbits.

The orbits bouncing off the wall from the focusing side are orbits with $|\phi| > \frac{1}{2} \pi$ in the phase space. In the phase space for sufficiently separated disks all parts of $\Lambda$ have angle $|\phi| < \frac{1}{2} \pi$. When $r$ decreases, some points in $\Lambda$ move closer to the lines $|\phi| = \frac{1}{2} \pi$ and the pruning starts at $r_c$ when the points in $\Lambda$ with the largest value of $|\phi|$ reach these lines.

When $r \rightarrow r_c$ with $r > r_c$, the value of $T$ in condition 1 tends to $\infty$. For $r < r_c$ the different rectangles of $\Lambda_t$ always have some overlapping. This overlapping corresponds to forbidden regions in the symbol plane. The forbidden regions that are the simplest to describe are the two regions consisting of points in $\Lambda$ with $|\phi| > \frac{1}{2} \pi$. The borders of
these regions in phase space are the lines $|\phi| = \frac{1}{2} \pi$. In the symbol plane the two regions correspond to two areas, one in the upper left-hand corner and the other in the bottom right-hand corner. The border is the symbolic representation of all points in $\Lambda$ with $\phi = \frac{1}{2} \pi$. To obtain this border, we scan through $x$ values with the angle $\phi = \pm \frac{1}{2} \pi$ and if $r > 2$ (an open system) in numerical work we only keep points bouncing more than 20 bounces both in the future and in the past.

As both $y$ and $\delta$ increase with the value of $x$, the border is monotonously increasing in the symbol plane.

The forbidden regions described above contain only forbidden orbits bouncing from the focusing side and not orbits passing through a forbidden region. We choose to take the pre-image of this described region as our primary forbidden region and the pre-image of the border as our pruning front. The other family of forbidden orbits, that is orbits passing through a disk, is an other primary forbidden region and has a second pruning front that is very similar to the first pruning front. An orbit on this second pruning front has the same symbol string as an orbit on the first pruning front but without the one symbol for the tangential bounce. The two primary forbidden regions have a straight line as a common border and this line corresponds to a gap between two bands in $A$ for sufficiently separated disks. We consider the two primary forbidden regions to be bounded by one pruning front. This region can be described as the primary overlapping in $\Lambda$. In one of the examples we will show that there may be more than one of these primary forbidden regions, but all are constructed in the same way.

The iteration of the points in the symbol plane is a shift operation of the symbols. This shift operation is slightly more complicated for well-ordered symbols than for the original disk symbols; see [23]. The primary forbidden region can be mapped forward and backward in time by the shift operation. The union of all images and pre-images of the primary forbidden region is dense and takes the full measure in the symbol plane. Their complement, the union of the legal orbits, has measure zero and is a Cantor set in the symbol plane.

From the primary forbidden region we can read off all the finite strings that can make an orbit forbidden. If the well-ordered alphabet has $M$ symbols, we divide the symbol plane into rectangles with side length $M^{-k}$. Each rectangle has a unique labelling $w_{-k+1}w_{-k+2}\ldots w_{k-1}w_k$ and corresponds to one disjoint area in the phase space of a sufficiently separated disk system. If the rectangle is inside the forbidden region, the symbol string $w_{-k+1}w_{-k+2}\ldots w_{k-1}w_k$ is forbidden in the alphabet. By excluding all the forbidden symbol strings for a given number $k$ we can redefine the symbolic dynamics in terms of a new alphabet with a new grammar, or in terms of a finite Markov graph [9, 10]. This is an approximation of order $k$ to the correct complete alphabet describing the billiard, and this gives an approximation to a Markov partition of the billiard. The implementation of such approximations in practice will be treated elsewhere.

3.1. Pruning for the three-disk billiard

The pruning in the three-disk billiard starts when the points in $\Lambda$ with the largest value of $|\phi|$ reach the value $|\phi| = \frac{1}{2} \pi$. These outermost points in $\Lambda$ are the heteroclinic points created by the crossing of the unstable (stable) manifold of the fixed point $W = \bar{0}$ and the stable (unstable) manifold of the fixed point $W = \bar{1}$. The numerical value of the critical distance is $r_c = 2.04821419\ldots$. The pruning front for three disks with $r < r_c$ is numerically determined by choosing points on disk 2 and starting the orbit by glancing off the disk at
point $x$ with the angle $\phi = \pm \frac{1}{2}\pi$. The first bounce in one direction is off disk 1 and in the other direction the bounce is off disk 3. Let $s_0 = 1$, and a point on the first pruning front is determined by choosing $s_1 = 2$ and $s_2 = 3$. The other values of $s_t$ are obtained from the numerical bouncing with starting point $x$ on disk 2. To find a point on the second pruning front we do not include the bounce in disk 2 but assume that the orbit grazes this disk. This gives the symbols $\tilde{s}_t = s_t$ for $t \leq 0$ and $\tilde{s}_t = s_{t+1}$ for $t > 0$. The points $(\gamma, \delta)$ are computed from the symbol strings of orbits with starting point $x$ between $\frac{3}{2}\pi$ and $\pi$. Plotting the points $(\gamma, \delta)$ in the symbol plane gives figure 7 for parameter $r = 2$ (touching disks). The pruning front is a monotonously increasing line from $(0.375, 0)$ to $(0.5, 0.5)$ and a monotonously decreasing line from $(0.5, 0.5)$ to $(0.75, 0)$.

A numerical test of this pruning front can be performed by starting one orbit at a random point in the billiard, letting it bounce in the billiard for some time and calculating the point $(\gamma, \delta)$ for each bounce. In figure 6(a) each bounce for an orbit with $10^6$ bounces is marked by a point in the symbol plane. As expected, the primary forbidden region is white in figure 6(a) as there are no bounces with these $(\gamma, \delta)$ values. In addition, there are three copies of this white region; the points $(\delta, \gamma)$ that are a result of the time reversion symmetry $(t \rightarrow -t)$ and the points $(1 - \gamma, 1 - \delta)$ and $(1 - \delta, 1 - \gamma)$ that are a consequence of the symmetry between odd and even in algorithm (1). This last symmetry is the result of the fact that our symbol plane includes two fundamental domains. All other white areas in figure 6(a) are images or pre-images of one primary forbidden region, where the iteration is a shift operation on the symbol string.

Figures 7 and 6(b) are magnifications of the symbol plane showing the pruning front and the distribution of points of a long orbit, respectively. The long chaotic orbit visits points close to both pruning fronts but the probability is different, as visualized in figure 6(b). The orbit is often close to grazing a disk but very seldom does the orbit bounce with angle close to $\frac{1}{2}\pi$. The tangential bounce is very unstable and this gives a low probability density to these bounces in the symbol plane. The stability is, however, not sensitive to how close an
orbit is to grazing a wall, and the probability density in the symbol plane approaches the second pruning front rather uniformly.

One can convert the topological pruning front in figure 7 into a symbolic description in different ways. The following method overcounts the number of allowed orbits.

In figure 7 rectangles with side length $2^{-7}$ are plotted together with the pruning front. A rectangle with side length $2^{-k}$ in the symbol plane corresponds to a symbol string of length $2k$ in the symbolic description. Each rectangle in figure 7 is identified with a symbol string $w_{-6}w_{-5}...w_{5}w_{7}$. If a rectangle is completely inside the primary forbidden region then the corresponding symbol string is forbidden. The length 14 symbol strings $w_{-6}w_{-5}...w_{5}w_{7}$ that are forbidden for three touching disks are listed in column 2 in table 2. Finding the completely forbidden strings of length $2k$ for all $k \in \{1,2,...\}$ gives a complete list of forbidden orbits. Table 2 gives this list for $k \leq 9$ when $r = 2$. The shortest forbidden symbol string is of length 12 ($k = 6$). In addition to the symbol strings $w$ in table 2 the $w'$ strings (obtained by letting $0 \rightarrow 1$ and $1 \rightarrow 0$) and the time reversed strings of $w$ and $w'$ have to be included as forbidden orbits.

It is possible to include in the table of forbidden strings the finite symbol strings describing rectangles that are just partly inside the primary forbidden region. This gives an undercounting of the legal orbits. It is also possible to apply a combination of these two methods. Other strategies are to use some periodic orbits close to the pruning front as an approximation, or to use some number of points that are exactly on the pruning front. It is presently not clear which approach yields the most convergent sequence of approximate grammars.

The pruning front is not a continuous line in the symbol plane. The images and pre-
Table 2. The forbidden orbits in three discs, \( r = 2 \). 'Under counting' approximation to level 9.

| \( W_5 \) ... \( W_6 \) | \( W_6 \) ... \( W_7 \) | \( W_7 \) ... \( W_8 \) |
|----------------|----------------|----------------|
| 000000 100000 | 000000 100001 | 000000 100010 |
| 000000 100000 | 000000 100010 | 000000 100100 |
| 010000 100000 | 010000 100010 | 010000 100100 |
| 010000 100000 | 110000 100000 | 010000 100010 |
| 001000 100000 | 110000 100000 | 110000 100010 |
| 000000 0111111 | 010000 100001 | 001000 100001 |
| 010000 0111111 | 010000 100000 | 010000 100000 |
| 110000 100000 | 110000 100000 | 110000 100000 |
| 000000 100000 | 000000 100000 | 000000 100000 |
| 001000 100000 | 001000 100000 | 001000 100000 |
| 010000 100000 | 010000 100000 | 010000 100000 |
| 011000 100000 | 011000 100000 | 011000 100000 |
| 000000 0111111 | 001000 100000 | 001000 100000 |
| 010000 0111111 | 010000 100000 | 010000 100000 |
| 011000 0111111 | 011000 100000 | 011000 100000 |
| 110000 0111111 | 110000 100000 | 110000 100000 |
| 000000 0111111 | 001000 100000 | 001000 100000 |
| 010000 0111111 | 010000 100000 | 010000 100000 |
| 011000 0111111 | 011000 100000 | 011000 100000 |
| 110000 0111111 | 110000 100000 | 110000 100000 |
| 000000 0111111 | 001000 100000 | 001000 100000 |
| 010000 0111111 | 010000 100000 | 010000 100000 |
| 011000 0111111 | 011000 100000 | 011000 100000 |
| 110000 0111111 | 110000 100000 | 110000 100000 |
| 000000 0111111 | 001000 100000 | 001000 100000 |
| 010000 0111111 | 010000 100000 | 010000 100000 |
| 011000 0111111 | 011000 100000 | 011000 100000 |
| 110000 0111111 | 110000 100000 | 110000 100000 |
| 000000 0111111 | 001000 100000 | 001000 100000 |
| 010000 0111111 | 010000 100000 | 010000 100000 |
| 011000 0111111 | 011000 100000 | 011000 100000 |
images of the pruning front cross the primary forbidden region and create open intervals in the pruning front. The result is that the pruning front itself is a Cantor set. The open intervals do not lead to any problems in converting the pruning front into a table of forbidden orbits. Any monotone curve in the interval, for example a straight line, gives a correct result.

The pruning front for an open three-disk system with \( r = 2.02 \) is plotted in figure 8 and that for a billiard with overlapping disks with \( r = 1.97 \) in figure 9. As expected, the figures show that the primary pruned region increases when the distance \( r \) decreases. The structure of the pruning front does not change dramatically when the distance changes.

![Figure 8. The pruning front for three disks with \( r = 2.02 \).](image)

![Figure 9. The pruning front for three disks with \( r = 1.97 \).](image)

3.2. Pruning for the four-disk billiard

The critical parameter value where the pruning starts for the four-disk system is \( r_c = 2.20469453 \ldots \). The pruning front is similar to the pruning front for the three-disk billiard. This billiard and the hyperbola billiard are discussed in detail in [24].

3.3. Pruning for a \((6 + 1)\)-disk system

The billiard with six disks on a large circle and one disk in the centre has pruned orbits when the distance between the disks is less than \( r_c = 3.59148407 \ldots \). Algorithm (3) with \( N = 6 \) gives the well-ordered symbols for this system. Figure 2 shows the seven disks and a part of one orbit that is grazing disk 7 (the centre disk) and the symbolic coordinate \((y, \delta)\) of this orbit is on the pruning front created by this centre disk. The symbol plane chosen here corresponds to the phase space of disk 1. In this system, disk 7 has a different phase space and symbol plane and the pruning front obtained is only valid for one of the disks on the large circle. There are two pruning fronts in the symbol plane; an orbit bouncing off disk 1 can graze either disk 7, as in figure 2 or it can graze disk 2. The primary pruned region in the symbol plane is limited by these pruning fronts, as shown in figure 10 for parameter \( r = 2.2 \). One orbit grazes both disk 2 and disk 7 and this orbit gives a point connecting the two pruning fronts.

The primary pruned region is converted to a list of forbidden symbol strings in a similar way to the three-disk billiard. Rectangles of size \( 5^{-k} \) correspond to symbol strings
We have shown that it is possible to define a well-ordered covering alphabet for different concave billiard systems with a finite number of concave walls. For these systems the primary forbidden region is a simple connected region in a symbol plane. The border of this region is the pruning front, and we have found this pruning front numerically for the three-disk system with different parameter values. The pruning front has also been obtained for a more complicated billiard with seven disks. The general method of constructing a symbolic alphabet ordered the same way as the ordering of orbits in the phase space has been
discussed. A well-ordered alphabet is necessary when finding the pruning front because, for any other alphabet, the primary pruned region is a very complicated region.

We show how the topological information, the pruned region, can be converted into symbolic dynamics, a list of forbidden symbol strings. The tables list some of these forbidden strings.

The method can easily be applied to other finite concave billiards with similar symmetries to those considered in this article. Infinite billiards like the Bunimovich stadium billiard, the Sinai billiard and the Lorentz gas systems are more complicated because it is difficult to obtain a well-ordered covering alphabet which is not pruned. These systems will be treated elsewhere.

A smooth potential described by the same symbolic dynamics as a disk billiard system [30] may have a pruning front and will also be treated in a similar way. There are, however, problems in defining a bounce in the smooth potential and this is presently under investigation.

Acknowledgments

The author is grateful to P Cvitanović, T Schreiber, F Christiansen, P Dahlqvist, M Sieber and Z Kaufmann for discussions. The author thanks the Norwegian Research Council for support.

References

[1] Cvitanović P, Percival I and Wirzba A (eds) 1992 Quantum Chaos—Quantum Measurement (Dordrecht: Kluwer)
[2] Cvitanović P (ed) 1992 Periodic Orbit Theory: theme issue, CHAOS 2 1–158
[3] Gaspard P and Rice S A 1989 J. Chem. Phys. 90 2225, 2242, 2255
[4] Sieber M and Steiner F 1990 Physica D 44 248; 1990 Phys. Lett. 148A 415
[5] Gutierrez M C 1990 Chaos in Classical and Quantum Mechanics (Berlin: Springer)
[6] Metropolis N, Stein M L and Stein P R 1973 J. Comb. Theory A 15 25
[7] Milnor J and Thurston W 1988 On iterated maps of the interval Dynamical Systems, Proc. University of Maryland 1986–87 (Lecture Notes in Mathematics 1342) ed A Dold and B Eckmann (Berlin: Springer) p 465
[8] Cvitanović P, Gunaratne G H and Procaccia I 1988 Phys. Rev. A 38 1503
[9] Cvitanović P 1989 Noise and Chaos in Nonlinear Dynamical Systems ed E Moss (Cambridge: Cambridge University Press); 1990 Nonlinear Physical Phenomena, Brasilia 1989 Winter School ed Â Ferraz, F Oliveira and R Osorio (Singapore: World Scientific); 1991 Applications of Chaos, EPRI workshop ed J H Kim and J Stringer (New York: Wiley)
[10] Grassberger P 1986 Int. J. Theor. Phys 25 907
[11] Cvitanović P private communication
[12] Hénon M 1976 Commun. Math. Phys. 50 69
[13] Isola S and Politi A 1990 J. Stat. Phys. 61 259
[14] D’Alessandro G, Grassberger P, Isola S and Politi A 1990 J. Phys. A: Math. Gen. 23 5285
[15] Cvitanović P 1991 Physica D 51 138
[16] Grassberger P and Kantz H 1985 Phys. Lett. 113A 235
[17] Eckhardt B and Wintgen D 1990 J. Phys. B: At. Mol. Opt. Phys. 23 355
[18] Christiansen F 1989 Masters Thesis University of Copenhagen
[19] Cvitanović P and Eckhardt B 1989 Phys. Rev. Lett. 63 823
[20] Cvitanović P and Eckhardt B 1993 Nonlinearity 6 277
[21] Dahlqvist P and Russberg G 1990 Phys. Rev. Lett. 65 2837
[22] Dahlqvist P and Russberg G 1991 J. Phys. A: Math. Gen. 24 4763
[23] Troy G 1991 Physica D 50 276
Symbolic dynamics I

[24] Hansen K T 1992 Chaos J. 2 71
[25] Smale S 1967 Bull. Am. Math. Soc. 73 747
[26] Lorentz H A 1905 Proc. Amst. Acad. 7 438
[27] Sinai Ya G 1970 Usp. Mat. Nauk 25 141
[28] Meiss J D 1992 CHAOS 2 267
[29] Hansen K T 1993 Symbolic dynamics: III. The stadium billiard, submitted to Nonlinearity
[30] Ezra G S, Richter K, Tanner G and Wintgen D 1991 J. Phys. B: At. Mol. Opt. Phys. 24 L413
   Wintgen D, Richter K and Tanner G 1992 Chaos J. 2