Colouring a Lorentz Gas

by

Steven Ole Nielsen

A thesis submitted in conformity with the requirements for the degree of Masters of Science Graduate Department of Chemistry University of Toronto

©Copyright by Steven Ole Nielsen 1998
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
To my parents and my sister
ABSTRACT

A colour label is introduced into a two-dimensional Lorentz gas system of three disks arranged in an equilateral triangle to represent chemical reactions. The microscopic level is studied systematically using an N-cylinder description where the exact dynamics is replaced by a symbolic dynamics which is a generating partition. The Kolmogorov-Sinai entropy and its finite and coloured varieties are discussed. These are then related to the coloured escape rate, a macroscopic property. Lastly, escape is eliminated by extending the three disk system to an infinite lattice, and the colour correlation function is explored. For large colouring regions the Poisson process rate law expression breaks down.
ACKNOWLEDGMENTS

I would like to thank my supervisor, Raymond Kapral, for suggesting an interesting problem and for discussions along the way. This work was supported by an NSERC postgraduate scholarship.
CONTENTS

Abstract ........................................ iii

Acknowledgments ................................ iv

Contents ........................................... v

List of Figures ................................. viii

1. Introduction ........................................ 1
   1.1 Description of system ................. 3
   1.2 Trapped trajectories ................. 4
   1.3 Definition of the repeller .......... 7
   1.4 Definition of $\theta$ and $\phi$ ...... 8
   1.5 Symbolic dynamics and the generating partition 9
   1.6 Definition of the N-cylinders ...... 10
       1.6.1 Discrete time version: a mapping 11

2. N-cylinders ........................................ 12
   2.1 Analytical results ................. 12
   2.2 Numerical results .................. 22
   2.3 Symmetry .................................. 29
9. Colour Correlation Function ........................................... 66
  9.1 Definition and a modification of the system ..................... 66
  9.2 Small colouring region behaviour ................................ 70
  9.3 A simple model ..................................................... 73
  9.4 Large colouring region behaviour .................................. 78

10. Conclusion .............................................................. 81

Appendix .............................................................................. 83
  A. Use of Probability .......................................................... 84
  B. Accessibility condition .................................................... 85
  C. Generalization for arbitrary R .......................................... 88
  D. Comments on numerical work ............................................ 91
  E. Additional details of Chapter 2.1 ....................................... 93
     E.1 Comparison with [1] .................................................. 93
     E.2 Explicit derivation of $\phi_{-}^{1-2}$ and $\phi_{+}^{1-2}$ .............. 95

Bibliography ........................................................................... 98
LIST OF FIGURES

1.1 Description of system .......................................................... 3
1.2 The inset shows the number of particles present in the system as a function of time. The main plot shows the same data with a logarithmic scale for the abscissa, which is in arb. units. The slope has a numerical value of $-\gamma = -0.662.$ ........................................... 5
1.3 Period 3 trapped trajectory ..................................................... 6
1.4 Definition of $\theta$ and $\phi$ ..................................................... 8
2.1 Geometric considerations for $\phi_{-12}^-$ .................................... 13
2.2 Geometric considerations for $\phi_{+12}^+$ ..................................... 14
2.3 Geometric considerations for obtaining $\phi_{-13}^-$ from $\phi_{-12}^-$. ... 16
2.4 Relation between $\phi^{in}$ and $\phi^{out}$ ....................................... 17
2.5 Geometric considerations for obtaining $\phi_{12}^-$ from $\phi_{12}^-$ ....... 19
2.6 Geometric considerations for obtaining $\phi_{32}^-$ from $\phi_{12}^-$ ........ 20
2.7 The upper left plot shows Cells 2.1 and 3.1. The upper right plot shows Cells 1.2 and 1.3. The lower plot shows cells 3.12, 2.13, 2.12, and 3.13 obtained by intersecting the upper plots. The dot in, for example, cell 3.12 means that we are looking at the trajectories while they are making their second collision with a disk. .................................................. 21
2.8 The 3-cylinder ‘123’.

2.9 The 3-cylinders broken up by hit number (disk 1 phase space only). The upper left plot shows hit #1, the upper right plot shows hit #3, and the lower plot shows hit #2.

2.10 The $\theta$ profile of the 3-cylinders. The upper left plot shows the profile of hit #1, the upper right plot shows the profile of hit #3, the lower left plot shows the profile of hit #2, and the lower right plot shows the net profile of all three hits. Each individual hit profile curve integrates to the same value. Abscissa in arb. units.

2.11 The 4-cylinders broken up by hit number. The upper left plot shows hit #1, the upper right plot shows hit #4, the lower left plot shows hit #2, and the lower right plot shows hit #3.

2.12 The 7-cylinders broken up by hit number. Reading left to right and top to bottom, we have hit numbers 1, 7, 2, 6, 3, 5, 4.

2.13 The level $N = 9$ approximation to the repeller, showing the density of the repeller as a function of $\theta$ (for disk 1 only by symmetry). Abscissa in arb. units.

2.14 The first collision frequency profile as a function of $\theta$. The $\theta$ values $\pi/6 - \pi/3$ are the reflection of $\theta = 0 - \pi/6$ in the line $\theta = \pi/6$. Abscissa in arb. units.

3.1 Effect of a coloured region.

5.1 Calculation of initial rate.
5.2 The coloured region spans $\theta \in [\pi/6 \pm 0.02]$ on disk 1. Reading left to right, top to bottom, snapshots of the black particle density are shown at times of $t = 0.35, 1.00, 1.50, 2.00, 2.60, 3.50, 10.00$. The remaining plot shows the black density as a function of time. After $t \approx 7$ the decay is roughly exponential. Note the initial linear segment.

5.3 The coloured region spans $\theta \in [0.02 \pm 0.02]$ on disk 1. Reading left to right, top to bottom, snapshots of the black particle density are shown at times of $t = 0.35, 1.00, 1.50, 2.00, 2.60, 3.50, 10.00$. The remaining plot shows the black density as a function of time. After $t \approx 7$ the decay is roughly exponential.

5.4 The upper curve corresponds to fig. 5.2. The lower curve corresponds to fig. 5.3. The middle curve is the phenomenological picture given by (4.1.19).

7.1 The negative of the coloured $h_{KS}^N$ for $N = 5$ (diamonds), $N = 7$ (crosses), and $N = 9$ (squares). Each symbol is the coloured Kolmogorov-Sinai entropy for a colouring region on disk 1 centered at the $\theta$-location of the symbol and extending one symbol to either side. Only $\theta = [0, \pi/6]$ is shown because $\theta = [\pi/6, \pi/3]$ is the reflection of this in the line $\theta = \pi/6$ (by symmetry).
7.2 Fig. 7.1 along with an approximation \((N = 9)\) to the repeller (which is also shown in fig. 2.13). The \(h_{KS}\) plots are to scale. Abscissa label omitted. .......................... 52

7.3 Construction of a modified Cantor set on the interval \([0, \pi/3]\). 53

7.4 Model \(h_{KS}^N\) with the equivalent of \(N = 5\) for the full system. 56

7.5 Fig. 7.1 superimposed on the negative of fig. 7.4. The discrepancy is thought to be mainly due to the one-dimensional nature of the model. .......................... 57

8.1 Negative of the black escape rate. Each diamond is the slope of the logarithm of the number of black particles in the system as a function of time. Each diamond represents a colouring region on disk 1 centered on the \(\theta\) location of the diamond and extending one diamond to either side. \(\theta \in [0, \pi/6]\) only shown because of symmetry. The escape rate for the uncoloured system (fig. 1.2) is 0.662, and this is also the value for the white escape rate for all of the data shown. .......................... 63

8.2 The logarithm of the number of black particles in the system at time \(t\) for colouring regions centered at \(\theta = 0.035\) (left plot) and \(\theta = 0.291\) (right plot) on disk 1. These correspond to the 4th and 33rd symbol from the left in fig. 8.1 respectively. Abscissa in arb. units. .......................... 64

8.3 Fig. 8.1 (black escape rate) and the \(N = 7\) curve of fig. 7.1 (coloured \(h_{KS}^7\)) superimposed to scale. Fig. 2.13 (the repeller density as a function of theta) is also shown, but not to scale. Abscissa label omitted. .......................... 65
9.1 Periodic extension of the three disk system. 

9.2 Appropriate boundary conditions to revert back to a three disk domain. A coloured region and three escaping/reinjected pairs of trajectories are shown. Each pair is coded differently (circle, square, blank) to clearly identify them.

9.3 The upper left plot shows the correlation function for a colouring region on disk 1 spanned by \( \theta \in [\pi/6 \pm 0.02] \). It also shows the correlation function obtained by sampling from the distribution (upper left plot of fig. 9.4) alone. The middle left plot shows the same curves, but for a colouring region covering all of disk 1. The upper curve is the actual correlation function, while the lower curve is the correlation function obtained by sampling from the distribution alone. The lower left plot shows the same curves again, but for a colouring region covering all three disks. The matching plots on the right hand side show the logarithms of these curves.

9.4 The distribution of times between colour changes. The upper left plot is for a colouring region on disk 1 spanned by \( \theta \in [\pi/6 \pm 0.02] \), the upper right plot is for the case where all of disk 1 is coloured, and the lower left plot is for the case where all three disks are completely coloured. The lower right plot shows the early time blowup of these distributions (drawing attention to their initial zero value). Abscissa in arb. units.

9.5 Physical picture for the constant time block correlation model.

The coloured region is shown as striped.
9.6 Considerations to compute the correlation function for \( t \in [0, T] \) in the constant time block model. ........................................ 75
9.7 Considerations to compute the correlation function for \( t \in [(i - 1)T, iT] \) in the constant time block model. ..................... 76
9.8 Physical picture for the variable time block correlation model.
The coloured region is shown as striped. ................................. 77
9.9 Considerations to compute the correlation function for \( t \in [0, T_{\text{min}}] \) in the variable time block model. ..................... 78
B.1 Violation of accessibility condition .............................................. 86
C.1 Generalized extrema of cell \(-12\) ............................................. 89
C.2 Geometric considerations for the generalized derivation of cell
\( 1 \cdot 2 \) from cell \(-12\) .............................................................. 90
E.1 Geometric considerations for \( \phi_{-2}^{1} \) ....................................... 94
E.2 Geometric considerations for \( \phi_{+2}^{1} \) ....................................... 96
1. INTRODUCTION

The justification of kinetic equations for systems involving chemical reactions is a fundamental problem of statistical mechanics. Most of the approaches involve phenomenological elements, and remain poorly justified from first principles. This is because, for systems within a classical mechanics framework, the microscopic description is deterministic and time-reversal symmetric. However, studies on chaos have shown that the dynamics of deterministic systems is very often unstable and hence that randomness is intrinsically generated by non-linear equations of time evolution. Chaos is characterized in terms of quantities defined at the microscopic level. One system that has received considerable attention in the literature is the Lorentz gas ([1–10]), which consists of a particle which moves freely between elastic collisions with fixed hard disk scatterers. It is known that this system is chaotic: it has a positive Lyapunov exponent and a positive Kolmogorov-Sinai entropy. In this thesis a colour label is introduced to represent chemical reactions with the aim of establishing links between the microscopic and macroscopic levels of description. To this end we focus on the Kolmogorov-Sinai entropy (and its finite and coloured varieties) and the global escape of coloured particles from the system. At the microscopic level we discretize the dynamics by imposing a symbolic dynamics on the system. This allows for a systematic
1. Introduction

approach by grouping the paths of particles (trajectories) according to how many disks they strike (these groupings are called N-cylinders) and leads to a simple expression for the Kolmogorov-Sinai entropy. At the macroscopic level an exact early time result is presented along with a late time picture. The chaotic nature of the system is observed to lead to a simple exponential law for the late time decay of trajectories due to escape. In this late time domain then, a phenomenological picture should be appropriate. The picture we adopt is shown in (1.0.1).

\[
\text{escape} \xleftrightarrow{\mathcal{B}} B \xrightarrow[k]{\mathcal{W}} \text{escape} \quad (1.0.1)
\]

Black and white particles interconvert and escape. The picture yields equations for the number of black and white particles as a function of time. An interesting question arises as to whether or not the black and white escape rates are different. This question is resolved and the resulting black escape rate data is compared to the coloured Kolmogorov-Sinai entropy. As a separate problem of interest, escape is eliminated by considering a system of infinite spatial extent. The colour correlation function is studied and is shown to be consistent with a Poisson process description for small colouring regions. For large colouring regions this description breaks down and hence we can say that a simple phenomenological description is not appropriate.

The outline of the thesis is as follows. Chapter 1 consists of an introduction and some definitions. In chapter 2 we look at the N-cylinders in detail. Colour is introduced in chapter 3. A phenomenological picture is given in chapter 4. The early time behaviour of the system is considered in chapter 5. Chapter 6 introduces the Kolmogorov-Sinai entropy; chapter 7 extends this concept to the coloured case. Chapters 6 through 4 are tied
together in chapter 8. Chapter 9 discusses the colour correlation function in a related system. Concluding remarks are made in chapter 10. Some additional material is presented in the appendices.

1.1 Description of system

![Diagram of a three disk Lorentz gas system]

Fig. 1.1: Description of system

We choose a three disk Lorentz gas system to work with. The centers of the three disks form an equilateral triangle (fig. 1.1). A point particle introduced into the system is subject to no forces other than elastic collisions with the curved walls and absorption by the straight walls. (This absorption will also
be referred to as 'escape', since if the straight walls were absent the particle would escape to infinity.) Hence the particle's speed remains at its initial value. This speed is chosen to be unity, so that time and distance traveled are equal in magnitude. An ensemble of (non-interacting) particles is used to calculate statistical quantities.

The value of $R = 3.11a$ is used for all of the numerical results. There are only two qualitatively different types of behaviour for the system: the value $R = 3a$ separates the system into 'tight' and 'loose' configurations (appendix B describes these two regions and shows that $R = 3a$ separates them). For the loose region the choice of $R = 3.11a$ is a compromise between a large $R$ value and computational feasibility. This study does not examine the tight configuration.

A generic particle will have a short lifetime unless the three disks are almost touching, effectively converting the open system to a closed one. It is empirically observed that the distribution of lifetimes follows roughly an exponential decay with some additional structure at short times (fig. 1.2).

1.2 Trapped trajectories

We know that some trajectories are trapped. For example, there is a trajectory which bounces repetitively between disks 1 and 2. Is there a period three trajectory? Yes, as is shown (fig. 1.3).

Consider the 3 points
Fig. 1.2: The inset shows the number of particles present in the system as a function of time. The main plot shows the same data with a logarithmic scale for the abscissa, which is in arb. units. The slope has a numerical value of $-\gamma = -0.662$. 
Fig. 1.3: Period 3 trapped trajectory
1. Introduction

(1.2.2) \( \left( a \cos \frac{\pi}{6}, a \sin \frac{\pi}{6} \right) \)

(1.2.3) \( \left( R + a \cos \frac{5\pi}{6}, a \sin \frac{5\pi}{6} \right) \)

(1.2.4) \( \left( \frac{R}{2} + a \cos \frac{3\pi}{2}, \frac{\sqrt{3}R}{2} + a \sin \frac{3\pi}{2} \right) \)

Start at (1.2.2) and go to (1.2.3). The normal to (1.2.3), which determines the reflection angle, makes an angle \( \pi/6 \) with the incoming trajectory, forcing the outgoing path to have an angle of \( 2\pi/3 \) measured in the usual way, \( \tan(2\pi/3) = -\sqrt{3} \), so that the line describing the outgoing path has equation \( y = -\sqrt{3}x + \sqrt{3}R - a \). Is it easily verified that (1.2.4) falls on this line. The outgoing path from (1.2.4) has slope \( \sqrt{3} \) since the normal at (1.2.4) is vertical. Hence the outgoing path follows the line \( y = \sqrt{3}x - a \). Point (1.2.2) is on this line. The incoming slope at (1.2.2) is \( \arctan(\sqrt{3}) = \pi/3 \), and the normal to (1.2.2) has slope \( \pi/6 \), so that the outgoing path has slope 0; back where we started.

1.3 Definition of the repeller

Due to the convex walls of the disks, the dynamics is everywhere defocusing. Hence there are no stable orbits. Yet some trajectories survive under the dynamics. The set of these form a fractal object called a repeller [11, p. 204]. Trajectories starting on the repeller will never leave it. This defines a probability distribution: the natural invariant probability distribution on a repeller tells us how often the various parts of the repeller are visited by trajectories that never escape [11, p. 207]. Trajectories coming close to the repeller are
responsible for the phenomenon of transient chaos, i.e., for chaotic behaviour that takes place for a long but finite time scale only [11, p. 204]. In order to study the trajectories systematically, we focus solely on the collision events, since the trajectories follow straight line paths in between.

1.4 Definition of $\theta$ and $\phi$

Two variables are needed to capture a collision: namely the position and velocity of the trajectory upon impact. We choose (following [1]) $\theta$ and $\phi$ as follows (fig. 1.4). The angle $\phi$ is measured from the outgoing trajectory to

![Diagram](image-url)

Fig. 1.4: Definition of $\theta$ and $\phi
1. **Introduction**

the normal. Hence for $\phi = \pi$ the trajectory is radial to the disk, while for $\phi = \pi \pm \pi/2$ the trajectory is tangential to the disk. The angle $\theta$ locates the point of impact on the disk and is measured in the usual manner, so that

\[
\begin{align*}
disk 1 & \text{ has } \theta \in [0, \pi/3], & disk 2 & \text{ has } \theta \in [2\pi/3, \pi], \\
& & \text{ and disk 3 has } \theta \in [4\pi/3, 5\pi/3].
\end{align*}
\]

(1.4.5)

1.5 **Symbolic dynamics and the generating partition**

Next, we introduce a symbolic dynamics and the notion of a generating partition. A partition of the phase space $\Gamma$ into $R$ subsets $A_i$ is defined by the two properties

\[
\begin{align*}
\bigcup_{i=1}^{R} A_i &= \Gamma, \\
A_i \cap A_j &= \emptyset.
\end{align*}
\]

(1.5.6)

(1.5.7)

Let us now follow a trajectory with initial condition $\zeta_0$ (which consists of both an initial position $x_0$ and an initial velocity $v_0$) and record its location at times $t_0 = 0, t_1, t_2, \ldots, t_n, \ldots$. The location of the particle $(x_0, x_1, x_2, \ldots, x_n, \ldots)$ at these instances falls into the subsets $A_{i_0}, A_{i_1}, A_{i_2}, \ldots, A_{i_n}, \ldots$. By this method we assign to each initial condition $\zeta_0$ and each time sequence $t_0 = 0, t_1, t_2, \ldots, t_n, \ldots$ a symbol sequence $i_0, i_1, i_2, \ldots, i_n, \ldots$. This is a course-grained description of the trajectory. The mapping from the phase space to the symbol space is called a symbolic dynamics. A given symbol sequence of finite length $N$ encompasses many different phase space sequences $\zeta_0, \zeta_1, \zeta_2, \ldots, \zeta_n$ because each symbol corresponds to a finite volume of phase space.
A generating partition is a partition where the infinite symbol sequence $i_0, i_1, i_2, \ldots, i_n, \ldots$ uniquely determines the initial condition $\zeta_0$ [11, pp. 35–36]. Thus the initial condition $\zeta_0$ determines the symbol sequence, and vice versa. A necessary condition for the existence of a generating partition is some sort of uniformity in the choice of the time sequences $t_0, t_1, t_2, \ldots, t_n, \ldots$.

In our case we choose the initial condition to be on the boundary of one of the disks, and choose time increments which match the collisions with disks. Hence our partition, which is a generating partition with this choice of time increments, consists merely of the boundaries of the three disks, and we denote it by

$$\{G\} = \{D_1, D_2, D_3\} \text{ or } \{1, 2, 3\} \text{ or } \{123\}.$$ (1.5.8)

Of course, to obtain an infinite symbol sequence the corresponding trajectory must remain in the system forever; it must be trapped and hence a member of the repeller. So for most initial conditions $\zeta_0$, only a finite symbol sequence will be generated and this particular symbol sequence contains many such trajectories as noted above. In addition, if the trajectory has its initial condition off the disk boundaries, then we will simply ignore its path until it hits a disk, at which point we record this collision as $i_0$.

### 1.6 Definition of the N-cylinders

We define the N-cylinder $(D_{i_1}, D_{i_2}, \ldots, D_{i_N})$ to be the set of all trajectories which hit disk $i_1$ on their first collision, disk $i_2$ on their second collision, \ldots, disk $i_N$ on their Nth collision, and then either go on to undergo further collisions or escape the system. The N-cylinders are useful both for an early-
time and a long-time description of the system because they partition the phase space into the number of hits, which corresponds roughly to the length of time a particle spends in the system. For early time phenomena one may look at, say, the complement of the 3-cylinders; for longer times at the 8-cylinders.

Each individual N-cylinder has contributions in 2 or 3 of the disk phase spaces, and there are $3 - 2^{N-1}$ total N-cylinders.

1.6.1 Discrete time version: a mapping

Talking in terms of cylinders and number of hits tends to reduce our thinking of the system to a mapping, where each time step corresponds to a particle progressing from one collision to the next. Indeed, the literature (and particularly textbooks) mainly treats deterministic chaos in terms of mappings. This avenue yields many results, and hence I will be guilty of talking about the continuous time system and its discretization into a mapping in the same breath.
2. N-CYLINDERS

2.1 Analytical results

The system changes in nature at \( R = 3a \). For \( R > 3a \) any point on one of the disks is accessible from any point on either of the other two disks by a direct path, provided all points are within their \( \theta \) bounds as given above (1.4.6). See appendix B for a proof of this.

The restriction of \( R > 3a \) applies to the following derivation. For a flavour of the approach required for \( R < 3a \) (and for more general systems) see appendix C. To begin with, we wish to find the trajectories which leave from disk 1 and hit disk 2, with both \( \theta \) variables in the range (1.4.6). Following[1], we call this region of phase space ‘cell \(-12\)’. The dot represents time. Specifically, the trajectories being considered are striking the disk which follows the dot, so that at this moment we may describe the trajectories in the \( \theta-\phi \) phase space of disk 1. This cell is bounded by two curves, which I will denote by \( \phi_{-12} \) and \( \phi_{+12} \). \( \phi_{-12} \) is the set of points mapped onto \( \theta_2 = \pi \) by the flow, while \( \phi_{+12} \) is the set of points mapped onto \( \theta_2 = 2\pi/3 \) by the flow.

The equation of the line from the generic point \((a \cos \theta, a \sin \theta)\) on disk 1 to the point \((R - a, 0)\) on disk 2 is (fig. 2.1)

\[
\frac{y - a \sin \theta}{x - a \cos \theta} = -\frac{a \sin \theta}{R - a - a \cos \theta},
\]  \hspace{1cm} (2.1.1)
Fig. 2.1: Geometric considerations for $\phi_{12}$
making the angle $\chi$ satisfy

$$
\tan \chi = -\frac{a \sin \theta}{a \cos \theta - R + a}.
$$

(2.1.2)

The angle we want is (fig. 2.1)

$$
\phi = \pi + \theta + \chi = \pi + \theta - \arctan \left( \frac{\sin \theta}{\cos \theta - R/a + 1} \right).
$$

(2.1.3)

See appendix E.1 for a comparison of this derivation with the result in [1].

Fig. 2.2: Geometric considerations for $\phi_{12}$
The equation of the line from the generic point \((a \cos \theta, a \sin \theta)\) on disk 1 to the point \((R + a \cos(2\pi/3), a \sin(2\pi/3))\) on disk 2 is (fig. 2.2)

\[
\frac{y - a \sin \theta}{x - a \cos \theta} = \frac{\sqrt{3}a/2 - a \sin \theta}{R - a/2 - a \cos \theta},
\]

making the angle \(\chi\) satisfy

\[
\tan \chi = \frac{\sqrt{3}a/2 - a \sin \theta}{R - a/2 - a \cos \theta}.
\]

The angle we want is (fig. 2.2)

\[
\phi = \pi + \theta - \chi = \pi + \theta - \arctan \left( \frac{\sqrt{3}/2 - \sin \theta}{R/a - 1/2 - \cos \theta} \right).
\]

\(\phi_{-13}\) is the set of points mapped onto \(\theta_2 = 5\pi/3\) by the flow, while \(\phi_{+13}\) is the set of points mapped onto \(\theta_2 = 4\pi/3\) by the flow. To obtain the curves \(\phi_{-13}\) and \(\phi_{+13}\) we appeal to symmetry, yielding (fig. 2.3)

\[
\phi_{+13}(\theta) = 2\pi - \phi_{-12}(\pi/3 - \theta), \quad (2.1.7)
\]
\[
\phi_{-13}(\theta) = 2\pi - \phi_{+12}(\pi/3 - \theta). \quad (2.1.8)
\]

Cell 1 · 2 is the image of cell ·12 under the dynamics. Specifically, we have allowed the trajectories to advance in time by one collision. It consists of the same trajectories, but now viewed within disk 2's phase space. The trajectories have struck disk 2 and are moving away from it. Hence the trajectory shown in fig. E.1 is an incoming, not an outgoing, trajectory. The implication is that the \(\phi\) shown in fig. E.1 is not the \(\phi\) we want. The
Fig. 2.3: Geometric considerations for obtaining $\phi^{13}$ from $\phi^{12}$
Fig. 2.4: Relation between $\phi^\text{in}$ and $\phi^\text{out}$
relationship uses the fact the incoming and outgoing trajectories make the same angle with the normal at the point of impact (fig. 2.4).

\[ \zeta = \phi^{\text{in}} - \pi \]  
\[ \phi^{\text{out}} = \phi^{\text{in}} - 2\zeta \]  
\[ \Rightarrow \phi^{\text{out}} = 2\pi - \phi^{\text{in}} \]  

These cells are obtained by symmetry arguments (fig. 2.5) It is also possible to derive these explicitly (see appendix E.2).

Note: the angle \( \phi \) show on disk 2 (fig. 2.5) is \( \phi^{\text{in}} \).

\[ \phi^{1:2}_-(\theta) = \phi^{1:2}_-(\pi - \theta) \]  
\[ \phi^{1:2}_+(\theta) = \phi^{1:2}_+(\pi - \theta) \]  

To get the cell \( 3 \cdot 2 \) we use symmetry (fig. 2.6)

\[ \phi^{3:2}_+(\theta) = 2\pi - \phi^{1:2}_+(5\pi/3 - \theta) \]  
\[ \phi^{3:2}_-(\theta) = 2\pi - \phi^{1:2}_-(5\pi/3 - \theta) \]  

In addition, the remaining cells of this form are given by displacing the curves to the remaining disk phase spaces. In this manner we get 12 cells (fig. 2.7, where only the phase space for disk 1 is shown). The intersections of these cells give us the middle hit of the 3-cylinders (fig. 2.7), and this gives an indication of how the hierarchy is constructed.
Fig. 2.5: Geometric considerations for obtaining $\phi^{1:2}$ from $\phi^{1:2}$
Fig. 2.6: Geometric considerations for obtaining $\phi^{3-2}$ from $\phi^{1-2}$
Fig. 2.7: The upper left plot shows Cells 2·1 and 3·1. The upper right plot shows Cells ·12 and ·13. The lower plot shows cells 3·12, 2·13, 2·12, and 3·13 obtained by intersecting the upper plots. The dot in. for example. cell 3·12 means that we are looking at the trajectories while they are making their second collision with a disk.
2.2 Numerical results

We now turn to numerical manifestations of the N-cylinders to explore them further. For the remainder of the thesis the dot representing time in the N-cylinder cells will be suppressed; its location will be clear from the context. For concreteness, let us look at the 3-cylinder ‘123’ (fig. 2.8). This cylinder consists of all trajectories which (from arbitrary initial conditions) first collide with disk 1, then disk 2, then disk 3. After this they may either escape or continue to collide with disks. We do not follow the trajectory after the 3rd hit. We wish to construct a ‘picture’ of this cylinder. We do this by plotting points (regions) in the aforementioned phase space. Each individual trajectory contributes 3 points, one for each collision. Let us now look at all 12 3-cylinders. Since the 3-disk system is symmetric in each disk, only the phase space corresponding to disk 1 will be shown. These cylinders are grouped by hit number as shown (fig. 2.9). These plots(fig. 2.9) may be simplified slightly by plotting \( \theta \) against intensity, where the intensity for a given \( \theta \) is a measure of how many phase space points fall into the vertical strip \((\theta - \epsilon, \theta + \epsilon)\) (fig. 2.10). It should be noted that the area under these plots is constant, since each trajectory contributes equally to each plot. Putting these three plots together gives the overall profile of the 3-cylinders (lower right plot in fig. 2.10). The 4-cylinders are shown in fig. 2.11, while the 7-cylinders are shown in fig. 2.12.

This gives us a good feel for the nesting structure of the N-cylinders, and of how the refinement proceeds with increasing N. A trajectory which hits N or more disks before exiting must of course first hit (N-1) disks. Furthermore, since trajectories belonging to an N-cylinder have no upper bound on the
Fig. 2.8: The 3-cylinder '123'
Fig. 2.9: The 3-cylinders broken up by hit number (disk 1 phase space only). The upper left plot shows hit #1, the upper right plot shows hit #3, and the lower plot shows hit #2.
Fig. 2.10: The $\theta$ profile of the 3-cylinders. The upper left plot shows the profile of hit #1, the upper right plot shows the profile of hit #3, the lower left plot shows the profile of hit #2, and the lower right plot shows the net profile of all three hits. Each individual hit profile curve integrates to the same value. Abscissa in arb. units.
Fig. 2.11: The 4-cylinders broken up by hit number. The upper left plot shows hit #1, the upper right plot shows hit #4, the lower left plot shows hit #2, and the lower right plot shows hit #3.
Fig. 2.12: The 7-cylinders broken up by hit number. Reading left to right and top to bottom, we have hit numbers 1, 7, 2, 6, 3, 5, 4.
number of collisions, this N-cylinder contains as subsets all higher cylinders (M-cylinders with M>N).

In the limit of N going to infinity, we are left with a fractal set of zero measure known as the repeller of the system (see section 1.3). This set consists of trajectories which are confined to the system for all time. Its structure should be apparent from the N-cylinder plots shown above. In these plots the data was segregated by hit number. Putting all of the hits together and focusing on the $\theta$ profile (like in the lower right plot in fig. 2.10), the level $N = 9$ approximation to the repeller is shown in fig. 2.13.

Fig. 2.13: The level $N = 9$ approximation to the repeller, showing the density of the repeller as a function of $\theta$ (for disk 1 only by symmetry). Abscissa in arb. units.
2.3 Symmetry

Let us explore the symmetry inherent in the progression of hit number distributions within a given N-cylinder. In general we notice that hit numbers \( i \) and \( N + 1 - i \) are similar (for example the 3 pairs of plots shown in fig. 2.12).

Consider the 3-cylinder '121'. As we know, this cylinder consists of all trajectories which, starting from arbitrary initial conditions, first strike disk 1, then disk 2, then disk 1, and then evolve naturally. The time period of interest, when we record events, is from the first to the third hit inclusive. Reversing time, we have trajectories which strike disk 1 from some initial distribution, then hit disk 2, then disk 1, and then evolve naturally. With the exception of a possible difference in the initial distribution of trajectories, these two scenarios are equivalent. But there is an asymmetry in the way \( \phi \) is measured. Since \( \phi \) is measured from the outgoing trajectory to the normal, and since reversing time changes incoming trajectories into outgoing trajectories we must use (2.1.11) to correct for this. Now consider the 3-cylinder '123'. The distributions of hit numbers 1 and 3 are not related to each other since the time-reversed version of '123' is '321', which is not equivalent to '123' under either of the cyclic permutations \( 1 \leftrightarrow 2, 2 \leftrightarrow 3, 3 \leftrightarrow 1; \quad 1 \leftrightarrow 3, 3 \leftrightarrow 2, 2 \leftrightarrow 1 \). So here hit 3 of '321' is related to hit 1 of '132', not hit 1 of '123'. But since we usually consider all 12 3-cylinders at once, every cylinder can be matched to its time-reversed partner.

We are left with the fact that the distributions are not perfectly equivalent, as seen in the upper plots of fig. 2.10. The explanation for this lies in the choice of initial conditions. Looking at the hit number distributions
for the 3-cylinders, it is observed that hit 3 has a flatter profile than hit 1. The profile for 1 is a consequence of the initial conditions, which are uniform random inside the 3-disk triangular domain. Due to these initial conditions, the first collision frequency profile as a function of $\theta$ is low at $\theta = 0, \pi/3$ and high at $\theta = \pi/6$ (fig. 2.14). This is because of the proximity of the extreme values of $\theta$ to the exit channels, from which no trajectories emanate.

![Graph](image)

*Fig. 2.14:* The first collision frequency profile as a function of $\theta$. The $\theta$ values $\pi/6-\pi/3$ are the reflection of $\theta = 0-\pi/6$ in the line $\theta = \pi/6$. Abscissa in arb. units.

Now consider hit 3. Firstly, every trajectory that gets plotted for hit 1 must be present for hit 3, by the definition of a 3-cylinder. Secondly, the dynamics is defocusing. Hence it is not surprising that the distribution at hit 3 is flatter. Hence I conclude that the difference in distributions is an artifact of the initial conditions.
3. COLOUR

3.1 Definition of colour

We introduce a binary label on the particles in the following way. Part of the boundary of the 3 disks is coloured. We are chiefly interested in the simplest case, whereby one continuous region is coloured. This colour can be thought of as a catalytic site, and its effect is to change the colour of particles which strike it (fig. 3.1). This can be thought of as a simple model of a chemical reaction. We could in principle have the colour change be a function of both $\theta$ and $\phi$, but we choose to accept all $\phi$, so that our 'reaction' does not depend on velocity.

Attaching a colour label does not change the motion of the particles, but it makes us focus on parts of the system that we would otherwise ignore or not distinguish between. Specifically, the colour label highlights a particular set of trajectories (eg those which change colour). Starting with an equal mixture of black and white particles defeats this purpose of distinguishing particular trajectories. In particular, the black and white escape rates will always be equal if we start with an equal mixture. Doing this obscures the label and is not useful. What makes much more sense is to start with all of the particles coloured white (or black, since the label is symmetric). By doing this the label serves its purpose most effectively.
Fig. 3.1: Effect of a coloured region
3.2 Location makes a difference

Let us now colour a small piece of disk 1. Then let us ask what difference it makes as to where the colouring region is located. Let us focus on the 3-cylinders. We group the 12 3-cylinders into a group of 8 and a group of 4. The division is made based on the disk number of the middle hit. The group of 4 consists of 212, 213, 312, and 313 (lower plot in fig. 2.9). The larger grouping (upper plots of fig. 2.9) is, to a first approximation, insusceptible to the location of the colouring region. However, the smaller grouping shows a very different behaviour. Specifically, for certain colouring choices no particles can change colour. Hence the coloured region might as well not be there in these cases. It is clear that the effect is caused by the particle’s hit number when it strikes a coloured disk, and that each hit number has its own measure over the disk. Hence we move away from talking about specific N-cylinders and instead talk about hit numbers.
4. PHENOMENOLOGICAL PICTURE – 2 ESCAPE RATES

The chaotic nature of the system is observed to lead to a simple exponential law for the late time decay of trajectories due to escape. In this late time domain then, a simple phenomenological picture should be appropriate. The picture we adopt is shown in (4.1.1). Black and white particles interconvert and escape. This picture yields equations for the number of black and white particles as a function of time \((n_B(t), n_w(t))\).

4.1 Derivation

\[
\begin{align*}
\text{escape} & \xleftrightarrow[k_B]{k_B} B \xrightarrow[k]{k} W \xrightarrow[k_w]{k_w} \text{escape} \\
\frac{dn_B}{dt} & = -kn_B + kn_w - k_B n_B \\
\frac{dn_w}{dt} & = -kn_w + kn_B - k_w n_w
\end{align*}
\]

(4.1.1)

(4.1.2)

(4.1.3)

We now find the eigenvalues of this system of equations.

\[
\begin{vmatrix}
-k - k_B - \lambda & k \\ k & -k - k_w - \lambda
\end{vmatrix} = 0
\]

(4.1.4)
\[
\lambda_{\pm} = -\frac{1}{2} \left\{ 2k + k_w + k_b \pm \sqrt{4k^2 + (k_w - k_b)^2} \right\}
\]  
(4.1.5)

To find the eigenvectors we solve the matrix equation

\[
\begin{bmatrix}
-k - k_b & k \\
k & -k - k_w
\end{bmatrix}
\begin{bmatrix}
\xi_{\pm}
\end{bmatrix}
= \lambda_{\pm} \xi_{\pm}.
\]  
(4.1.6)

This yields

\[
\xi_{\pm} = \left( \frac{\lambda_{\pm} + k + k_w}{k}, 1 \right).
\]  
(4.1.7)

Hence

\[
n_B(t) = \left[ \frac{k}{\sqrt{4k^2 + (k_w - k_b)^2}} \right] \left[ \xi_-(t) - \xi_+(t) \right]
\]  
(4.1.8)

\[
n_w(t) = \left[ \frac{k}{\sqrt{4k^2 + (k_w - k_b)^2}} \right] \left[ \left( \frac{\lambda_+ + k + k_w}{k} \right) \xi_-(t) 
- \left( \frac{\lambda_- + k + k_w}{k} \right) \xi_+(t) \right].
\]  
(4.1.9)

Substituting fully gives

\[
n_B(t) = \left[ \frac{k}{\sqrt{4k^2 + (k_w - k_b)^2}} \right] \left\{ \left( \frac{\lambda_- + k + k_w}{k} \right) n_B(0) + n_w(0) \right\} e^{\lambda_- t}
- \left\{ \left( \frac{\lambda_+ + k + k_w}{k} \right) n_B(0) + n_w(0) \right\} e^{\lambda_+ t},
\]  
(4.1.10)
4. Phenomenological Picture – 2 escape rates

\[ n_w(t) = \left[ \frac{-k}{\sqrt{4k^2 + (k_w - k_b)^2}} \right] \left[ \left( \frac{\lambda_+ + k + k_w}{k} \right) n_B(0) + n_w(0) \right] e^{-\lambda_- t} \]

\[ - \left( \frac{\lambda_+ + k + k_w}{k} \right) \left[ \left( \frac{\lambda_- + k + k_w}{k} \right) n_B(0) + n_w(0) \right] e^{\lambda_+ t} \].

If the two escape rates are in fact not distinct, then we get

\[ n_B(t) = \frac{1}{2} \{ [n_B(0) + n_w(0)] e^{-k_e t} - [-n_B(0) + n_w(0)] e^{(-2k_e) t} \}, \]

\[ n_w(t) = \frac{1}{2} \{ [n_B(0) + n_w(0)] e^{-k_e t} + [-n_B(0) + n_w(0)] e^{(-2k_e) t} \}. \]

These last two equations (4.1.12, 4.1.13) can be derived from scratch, or by setting \( k_b = k_w = k_e \) in (4.1.10) and (4.1.12). Note that

\[ |\lambda_-| < |\lambda_+|, \]

since

\[ |\lambda_+| - |\lambda_-| = \sqrt{4k^2 + (k_w - k_b)^2} > 0. \]

In chapter 8 we will see that the escape rates are distinct.

Can the coefficient of \( \exp \lambda_- t \) ever drop out, to give a different asymptotic slope? Clearly not in (4.1.12, 4.1.13). In (4.1.10, 4.1.12), we must check if the coefficients of \( n_B(0) \) and \( n_w(0) \) always have the same sign. The coefficient of
4. Phenomenological Picture – 2 escape rates

$n_B(0)$, up to a positive constant\(^1\), is $\lambda_- + k + k_w$. But

\[
\lambda_- + k + k_w = -\frac{1}{2} [2k + k_w + k_B - \sqrt{4k^2 + (k_w - k_B)^2}] + k + k_w
\]
\[
= \frac{1}{2} [k_w - k_B + \sqrt{4k^2 + (k_w - k_B)^2}] \\
\geq \frac{1}{2} [k_w - k_B + |k_w - k_B|] \\
\geq 0.
\] (4.1.16)

Specializing to the initial conditions

\[
n_B(0) = 0, \quad n_w(0) = 1.
\] (4.1.17)

the black density is given by

\[
n_B(t) = \frac{k}{\sqrt{4k^2 + (k_w - k_B)^2}} [e^{\lambda_- t} - e^{\lambda_+ t}]
\] (4.1.18)

when the escape rates are distinct, and by

\[
n_B(t) = \frac{1}{2} [e^{-k_w t} - e^{(-2k_w - k) t}]
\] (4.1.19)

when the escape rates are equal. The initial slope is obtained by expanding the exponential functions around $t = 0$ to linear order, and is $k$ for both (4.1.18) and (4.1.19). This initial slope is examined further in chapter 5.

The predicted asymptotic slope of $\lambda_-$ includes all three variables of (4.1.1) in the case where the escape rates are distinct. In chapter 8 we isolate the black escape rate $k_w$. For the small colouring regions considered, the white

\(^1\) What I really mean is that the constant I am ignoring is common to both terms (black & white), so that the relative sign between the two is unaffected.
escape rate does not change measurably from the uncoloured value (ie. it does not vary with the location of the colouring region). The much simpler single variable dependence of the asymptotic slope in (4.1.12,4.1.13) would give us a greater understanding of the system, but as we shall see in chapter 8 this is not the case. So our phenomenological picture is not so simple after all. In section 5.2 we will see that it does not describe the system at early times.
5. INITIAL EVENTS

5.1 Derivation of initial rate

Let us take a uniform initial velocity and position distribution, and an initial colour condition of all white (this initial condition was shown in (4.1.17)). The phenomenological picture of chapter 4 predicts an initial (constant) rate of production of black particles of \( k \). We wish to consider the exact early time dynamics in the presence of a coloured region. The coloured region is immediately subjected to a constant flux of white particles, and hence the initial production of black particles occurs at a constant rate. This rate is evaluated as follows.

Take a disk and chop its circumference into \( n \) equal pieces, so that each is subtended by an arc \( \theta = 2\pi/n \) (fig. 5.1). Choose the midpoint on the circumference of each of these regions and approximate the circumference by a straight line tangent to the disk at this point. The length of this line segment is precisely \( 2a \tan(\pi/n) \) which we will approximate as \( 2\pi a/n \).

We wish to look at trajectories hitting the line segment within time \( \Delta t \). Both position and velocity space must be taken into account.

Fixing the velocity to a specific direction \( \phi \), the allowed position space is
given by a parallelogram of area (shaded area in fig. 5.1)

\[ A = \frac{2\pi a}{n} \Delta t \sin \phi. \]  

(5.1.1)

Hence the total phase space contribution is

\[ 2 \int_0^{\pi/2} \frac{2\pi a}{n} \Delta t \sin \phi \, d\phi = \frac{4\pi a \Delta t}{n}. \]  

(5.1.2)

Now Riemann says, for the entire disk,

\[ \sum_{i=1}^{n} \frac{4\pi a \Delta t}{n} = 4\pi a \Delta t. \]  

(5.1.3)
This result is valid for any time $\Delta t$. Also it allows the area of interest to be a
coloured region instead of the entire disk. For a colouring region subtended
over an angle of $\varphi$ the volume of phase space hitting within time $\Delta t$ is $2\varphi a\Delta t$,
giving a constant rate of $2\varphi a$.

5.2 Snapshots and range of validity

This rate remains valid until the population of white particles has been de-
pleted sufficiently. This depletion shows up beautifully in the snapshots
below, where black particles only are shown (figs. 5.2, 5.3). Let us look at
these series in more detail.

In the first set (fig. 5.2) the coloured region is centered at $\theta = \pi/6$
on disk 1 and spans a total width of 0.04 radians. As time increases from
zero white particles strike the colored region, turning black and bouncing
outwards from this region. There are two points worth noting at $t = 1.00$.
Firstly, black particles have propagated to the boundary and are escaping the
system. Secondly, while at $t = 0.35$ the black particles which stay closest to
disk 1 follow a tangential path outwards from the coloured region, at $t = 1.00$
the shape of the white space between disk 1 and the black region is different.
To see why this is so, imagine a black particle somewhere in this white gap.
and then deduce where it would have had to start from (as a white particle)
at $t = 0$. Recall that the particles move at a speed of unity, so that the
path length from $t = 0$ to $t = 1.00$ must be 1 unit. This effectively reduces
the range of the integration limits in (5.1.2). At $t = 1.50$ some of the black
particles have struck disks 2 and 3. At $t = 2.00$ the fronts propagating from
disks 2 and 3 have almost met; parts of these fronts are visible in the white
gap region mentioned above. Most striking is the appearance of a hole close
the coloured region. By now all direct collisions with the coloured region have
occurred — all particles striking the coloured region must have first struck
disk 2 or disk 3. This constrains the velocity vectors of the newly coloured
black particles. At \( t = 2.60 \) this hole has grown, and the black particles
rebounding from disks 2 and 3 have reached disk 1, so that a few of the black
particles can now change colour. The middle hole has opened completely at
\( t = 3.50 \); this is the cause of the dip in the time profile plot. The loss of
all discernible structure at \( t = 10.00 \) is worth noting. The time profile plot
shows the number of black particles present in the system as a function of
time and corresponds to the snapshots.

In the second set (fig. 5.3), where the coloured region is at the edge
of the disk 1 boundary (and still spans a total width of 0.04 radians), the
integration limits in (5.1.2) are invalidated immediately. This is because
white particles do not initially surround the coloured region (none are below
it). This proximity to an exit channel results in a very different behaviour
from fig. 5.2. Black particles can escape the system immediately. Looking at
\( t = 0.35 \) and 1.00, we see that almost from \( t = 0 \) the white particles striking
the coloured region have a highly restricted set of velocity vectors. Black
particles collide with disks 2 and 3 at \( t = 1.50 \) and \( t = 2.00 \) respectively.
Recollisions with disk 1, which result in some black particles striking the
coloured region, occur at \( t = 2.60 \). After some time \( (t = 10.00) \) almost all
discernible structure has been lost. The time profile plot shows the number
of black particles present in the system as a function of time and corresponds
to the snapshots.
Fig. 5.2: The coloured region spans $\theta \in [\pi/6 \pm 0.02]$ on disk 1. Reading left to right, top to bottom, snapshots of the black particle density are shown at times of $t = 0.35, 1.00, 1.50, 2.00, 2.60, 3.50, 10.00$. The remaining plot shows the black density as a function of time. After $t \approx 7$ the decay is roughly exponential. Note the initial linear segment.
Fig. 5.3: The coloured region spans $\theta \in [0.02 \pm 0.02]$ on disk 1. Reading left to right, top to bottom, snapshots of the black particle density are shown at times of $t = 0.35, 1.00, 1.50, 2.00, 2.60, 3.50, 10.00$. The remaining plot shows the black density as a function of time. After $t \approx 7$ the decay is roughly exponential.
Let us examine the time profiles for the above two colouring choices in view of both section 5.1 and the phenomenological picture of chapter 4. We will work with (4.1.19) instead of (4.1.18) because (4.1.18) has too many parameters to fit. (Guessing at some of the quantities in (4.1.18) gives a very similar result to (4.1.18) for what follows.) The derivation in section 5.1 predicts an initial rate (with a colouring region of width 0.04 radians) of 0.08. This value agrees with fig. 5.2 as expected, while fig. 5.3 bends away from this value immediately. We need two parameters for (4.1.19), namely $k$, which was shown to be equal to the initial slope, and $k_e$, the asymptotic decay constant. Taking $k = 0.08$ and $k_e = 0.56$ (the observed decay constant of figs. 5.2 and 5.3), the phenomenological picture is shown in fig. 5.4 along with the time profiles from figs. 5.2 and 5.3. In the phenomenological picture escape and reaction occur all of the time at fixed rate. This is very different from what actually occurs. In fig. 5.2 there is an initial buildup of black density with no escape. This accounts for the difference in height between these two curves. In fig. 5.3 the opposite extreme occurs, where black density escapes more quickly than in the phenomenological picture. Hence our phenomenological picture is not valid in the early time regime where the dynamics is very structured.
Fig. 5.4: The upper curve corresponds to fig. 5.2. The lower curve corresponds to fig. 5.3. The middle curve is the phenomenological picture given by (4.1.19).
6. KOLMOGOROV-SINAI ENTROPY $H_{KS}$

Recall that by looking at the symbolic dynamics we have discretized our system – i.e., turned it into a mapping, where each ‘time step’ is from one collision to the next.

6.1 Definition

We denoted our generating partition by (1.5.8)

$$\{G\} = \{D_1, D_2, D_3\}. \quad (6.1.1)$$

Recall that the $N$-cylinder $(D_{i_1}, D_{i_2}, \ldots, D_{i_N})$ is the set of all trajectories which hit disk $i_1$ on their first collision, disk $i_2$ on their second collision, $\ldots$, disk $i_N$ on their $N$th collision, and then either go on to undergo further collisions or escape the system. The probability of observing a given $N$-cylinder $N_i$ is denoted by $p(i_1, i_2, \ldots, i_N)$. It depends upon the initial distribution $\sigma$.

The Shannon information, which plays an important role in dynamical systems theory, is given by

$$I(p) = \sum_{i_1, i_2, \ldots, i_N} p(i_1, i_2, \ldots, i_N) \ln p(i_1, i_2, \ldots, i_N). \quad (6.1.2)$$

The summation in (6.1.2) is over all allowed symbol sequences, i.e., over
events with positive probability. The Shannon entropy is then

\[ H = -I(p). \]  \hspace{1cm} (6.1.3)

This entropy depends on several quantities:

- the probability distribution \(\sigma\) of the initial values,
- the partition of the phase space used (here we used the generating partition \(\{G\}\)),
- the length \(N\) of the symbol sequence.

The Kolmogorov-Sinai entropy is defined as

\[ h_{KS} = -\sup_{\{\mathcal{P}\}} \lim_{N \to \infty} \frac{1}{N} \sum_{i_1, i_2, \ldots, i_N} p(i_1, i_2, \ldots, i_N) \ln p(i_1, i_2, \ldots, i_N). \]  \hspace{1cm} (6.1.4)

where a supremum is taken over all possible partitions of the phase space and where \(\sigma\), the measure of initial conditions, is taken to be \(\mu\), the natural invariant measure of the repeller.

I have only considering one partition, namely the generating partition \(\{G\}\), since it can be proved that this special partition is in fact the partition which yields the supremum [11, p. 148]. So we immediately have

\[ h_{KS} = -\lim_{N \to \infty} \frac{1}{N} \sum_{\{i_1, i_2, \ldots, i_N\} \in \{G\}} p(i_1, i_2, \ldots, i_N) \ln p(i_1, i_2, \ldots, i_N). \]  \hspace{1cm} (6.1.5)

Furthermore, at level \(N\), the logical approximation to \(\mu\) is simply the probability distribution associated with the \(N\)-cylinders; i.e., the level \(N\) coarse grained distribution.

A chaotic system can be defined as one in which \(h_{KS} > 0\). This quantity gives the data-accumulation rate necessary to follow the time evolution and to recover the trajectory of the system from the recorded data [2].
6.2 Maximum possible value

If every N-cylinder is of equal probability, then

$$p(i_1, i_2, \ldots, i_N) = (3 \cdot 2^{N-1})^{-1},$$  \hspace{1cm} (6.2.6)

so that at level N,

$$h^K_{KS} = -\frac{1}{N} \ln \left((3 \cdot 2^{N-1})^{-1}\right)$$

$$= \ln 2 + \frac{\ln 3 - \ln 2}{N}. \hspace{1cm} (6.2.7)$$

It is observed that for our system this formula (6.2.7) seems to hold approximately, but only low N values are feasible computationally (see table 6.1).

<table>
<thead>
<tr>
<th>N</th>
<th>$H^K_{KS}^N$</th>
<th>maximum (see 6.2.7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.828</td>
<td>0.8283</td>
</tr>
<tr>
<td>5</td>
<td>0.772</td>
<td>0.7742</td>
</tr>
<tr>
<td>7</td>
<td>0.748</td>
<td>0.7511</td>
</tr>
<tr>
<td>9</td>
<td>0.733 ± 0.002</td>
<td>0.7382</td>
</tr>
<tr>
<td>11</td>
<td>0.721 ± 0.01</td>
<td>0.7300</td>
</tr>
</tbody>
</table>

*Tab. 6.1: Uncoloured $H^K_{KS}^N$*
7. COLOURED $H_{KS}$

7.1 Definition

For a coloured version of $h_{KS}$ we assign probabilities for the various N-cylinders from the subset of trajectories which escape coloured black (from an all white initial condition). Fig. 7.1 shows some finite approximations of the coloured entropy for small colouring regions (see figure caption for details). There is a qualitative similarity with the $\theta$-profile of the repeller, as shown in fig 7.2. This resemblance is further explored in the following section.

7.2 Model

7.2.1 Motivation

This model is motivated by a desire to understand the structure in fig. 7.1. However, as we will see, the model is too simple to capture the behaviour of the system at large $N$. For a small enough colouring region the number of recollisions with the coloured region is negligible and the statistics of the hits are governed by the hit number distributions alone – i.e. there are no 'dynamical effects'. For this reason the model to be presented shortly has no dynamics. Since the colouring region is a function of $\theta$ only, the model focuses
Fig. 7.1: The negative of the coloured $h_{KS}^N$ for $N = 5$ (diamonds), $N = 7$ (crosses), and $N = 9$ (squares). Each symbol is the coloured Kolmogorov-Sinai entropy for a colouring region on disk 1 centered at the $\theta$-location of the symbol and extending one symbol to either side. Only $\theta = [0, \pi/6]$ is shown because $\theta = [\pi/6, \pi/3]$ is the reflection of this in the line $\theta = \pi/6$ (by symmetry).
Fig. 7.2: Fig. 7.1 along with an approximation \((N = 9)\) to the repeller (which is also shown in fig. 2.13). The \(h_{KS}\) plots are to scale. Abscissa label omitted.
on the projection of the hit number distributions presented earlier. These

distributions have almost identical projections for the pairs of hit numbers

\( i \) and \( (N + 1 - i) \) (eg. fig. 2.11). We could include this in the model by

making all of the model's levels doubly degenerate. However, this factor of

two drops out upon normalizing the probabilities, so we leave it out. One

useful aspect of the model is that it is possible to take the limit \( N \to \infty \) in

the Kolmogorov-Sinai entropy.

\subsection*{7.2.2 Construction of the model}

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{fig7.3}
\caption{Construction of a modified Cantor set on the interval \([0, \pi/3]\)}
\end{figure}
Consider a modified Cantor set as follows. The simple Cantor set is constructed from $[0, 1]$ by repeatedly removing the middle third of every interval present. The modified set is constructed from $[0, \pi/3]$ in the following manner. The first and last quarter of this interval are preserved, and a quarter centered on the middle of this interval is also kept. The remaining two segments (an eighth of the interval each) are discarded. The density of the outer segments remains the same as that of the interval they came from, while the density of the middle segment is chosen to be twice this. Another way to think of this operation is to imagine preserving the outer quarters while folding the rest of the interval onto the middle quarter (fig. 7.3). This operation is carried out on every interval, generating in the limit a fractal set.

A few comments are in order here. The total density is preserved at each stage. This is because for fixed $N$ in the full system, each hit number plot contains the same total density as discussed earlier. Looking at the hit number 1 and $N$ distributions for the full system (eg. upper plots in fig. 2.11), we see that in projection onto the $\theta$ axis all of the $N$-cylinders are degenerate and span the entire $\theta$ range. The next level of hit numbers (hits 2 and 3), in projection again, occupy three intervals and hence the $N$-cylinders are not fully degenerate. In fact from the lower plots in fig. 2.11 we see that half of the $N$-cylinders are degenerate in the middle interval, while one quarter are degenerate to either side. This pattern continues as we move towards the middle hit number. This accounts for the method of construction of the Cantor set. The gap chosen for the Cantor set ($1/4$ of the total interval lengths are discarded with each step) corresponds to a particular choice of $R$ in the full system. This is because changing $R$ in the full system changes
the size of the gaps (in projection onto the \( \theta \) axis) between intervals in, for example, the lower plots in fig. 2.11. The gap in the model was chosen more out of convenience (to keep the density constant) than with \( R = 3.11a \) in mind — it makes no qualitative difference.

The initial interval \([0, \pi/3]\) is denoted as the 'level 0' Cantor set. The next iteration is denoted as the 'level 1' Cantor set. In general, the set obtained after \( n \) refinements as described above is denoted as the 'level \( n \)' Cantor set. Adding up the total number of intervals in a particular level of refinement and weighting each interval by its density gives us \( 2^{2n} \) intervals for the 'level \( n \)' set. We choose a number \( N \) and a point \( \theta \in [0, \pi/3] \). We consider the first \( N \) levels of our Cantor set. Level \( N \) and consists of \( 3^N \) intervals where the degeneracy of an interval is given by its weight (density). For levels \( i \) above this \((i < N)\), the degeneracy of an interval is taken to be its weight multiplied by the factor \( 2^{2(N-i)} \). Doing this gives us a total of \( 2^{2N} \) (possibly degenerate) intervals for every level of the Cantor set, which we want since every hit number distribution for the full system contains the same total number of \( N \)-cylinders. For example level 0 consists of \( 2^{2N} \) degenerate intervals (all \([0, \pi/3]\)). Adopting the notation used throughout this work, we call the \( 2^{2N} \) intervals \( N \)-cylinders. For each level, we ask whether or not the point \( \theta \) is contained in some interval. If so, we add a probability contribution of \( p \) to each of the \( N \)-cylinders present at this point. Next, we calculate the value of \( p \) necessary to give an overall probability of 1 (ie. we normalize \( p \)). Finally, the level \( N \) Kolmogorov-Sinai entropy is calculated using

\[
h_{KS}^N = -\frac{1}{N} \sum_{i=1}^{2^N} p(i) \ln p(i),
\]

(7.2.1)
where \( p(i) \) is the probability present in the \( i^{th} \) N-cylinder. Fig 7.4 shows an example of this. The level \( N \) in the Cantor-type model corresponds to the level \((2N + 1)\) in the full system. This is because the full system has the double degeneracy mentioned above.

![Diagram of model](image)

*Fig. 7.4: Model \( h_{KS}^N \) with the equivalent of \( N = 5 \) for the full system*

The discrepancy in fig. 7.5 is thought to be mainly due to the one-dimensional nature of the model. The plot for the full system has a pronounced ‘baseline drift’ in comparison to the Cantor model. The model is a simplified representation of the full system since each interval (at each N-cylinder level) in projection is not in fact of constant density, as seen for example in the lower left plot of fig. 2.10 where, in projection, there are
Fig. 7.5: Fig. 7.1 superimposed on the negative of fig. 7.4. The discrepancy is thought to be mainly due to the one-dimensional nature of the model.
three regions of non-zero intensity.

7.3 Limit as $N \to \infty$

The assumption of trajectories striking the coloured region at most once clearly breaks down as $N$ grows. For instance, all periodic and quasi-periodic trajectories violate this condition. This assumption is the basis for using the $N$-cylinders (for a system without colour changing regions) as a static description of the probabilities present in the Kolmogorov-Sinai entropy expressions. The breakdown of this assumption invalidates this approach. This is because the recollision events are conditionally dependent on the prior collisions. This could be called a dynamical effect and it changes the (now coloured) $N$-cylinders in complicated ways. To give the simplest example of this effect, note that there is a minimum number of collisions needed ($\geq 2$) from one colour change to the next at the same coloured location. Hence the model loses its connection to the full system at large $N$. The infinite limit is presented anyways to be complete.

To show the range of values in the Cantor-type model, I will consider the two extreme cases, which correspond to the extremal values for $h_{KS}^N$. The first case is for $\theta \in (\pi/12, \pi/8) \cup (5\pi/24, \pi/4)$ and corresponds to the maximal $h_{KS}^N$ value. The second case is for $\theta = \pi/6$ and corresponds to the minimal $h_{KS}^N$ value.
7.3.1 Case 1: Maximum value

All $2^{2N}$ N-cylinders (bins) have an equal contribution of $p$ in this case. Hence $p = 2^{-2^N}$ and

$$h_{KS}^N = -\frac{1}{2N} \sum_{i=1}^{2^{2N}} p \ln p \quad (7.3.2)$$

$$\Rightarrow h_{KS} = \ln 2 \quad (7.3.3)$$

7.3.2 Case 2: Minimum value

Fix $N$ for now. Out of the $2^{2N}$ cylinders present at each level, $2^{2N}$ have a contribution of $p$ ($0^{th}$ level), $2^{2N-1}$ have an additional contribution of $2p$ ($1^{st}$ level), $2^{2N-2}$ of these have an additional contribution of $4p$, etc. Hence

$$1 = p \left(2^{2N} + 2^{2N-1}2 + \cdots + 2^{2N-N}2^N\right)$$

$$\Rightarrow p = \left(\left(N+1\right)2^{2N}\right)^{-1}, \quad (7.3.4)$$

Whence

$$h_{KS}^N = -\frac{1}{2N} \left[\left(2^{2N} - 2^{2N-1}\right) p \ln p + \left(2^{2N-1} - 2^{2N-2}\right) 3p \ln 3p + \right.$$\right.$$\left.$$\cdots + \left(2^{N+1} - 2^{N}\right) (2^N - 1)p \ln[(2^N - 1)p]+ \right.$$\right.$$2^N(2^{N+1} - 1)p \ln[(2^{N+1} - 1)p]$$

$$= -\frac{1}{2N} \sum_{j=1}^{N} \left(2^{2N-j+1} - 2^{2N-j}\right) (2^j - 1)p \ln[(2^j - 1)p] +$$

$$2^N(2^{N+1} - 1)p \ln[(2^{N+1} - 1)p]. \quad (7.3.5)$$

This sum can be performed analytically if all of the $(2^j - 1)$ terms are replaced with $2^j$ (using Maple; not shown). The limit as $N$ goes to infinity for this
substituted expression is \( \frac{3}{4} \ln 2 \). Since this replacement increases the value of the expression, this limit is an upper bound. It is numerically observed that the actual sum is very close to this substituted expression for \( N < 1500 \).

The model is qualitatively accurate for small \( N \) (and for the small colouring regions considered) because the assumption of trajectories striking the colouring region at most once during their lifetime is justified. This assumption is not valid at large \( N \) and invalidates the model. The \( \theta \)-profile density of the repeller is also qualitatively similar to the coloured Kolmogorov-Sinai entropy. This reflects the nature of the set of trajectories being considered in the entropy calculation for each colouring region. Specifically, the variations in entropy with the location of the colouring region show that the sets of trajectories underlying these entropies do not sample the entire phase space available to all of the trajectories (irrespective of colour), but are predominately sampling the local environment around the colouring region.
8. ESCAPE RATES & LINKS BETWEEN MICROSCOPIC AND MACROSCOPIC QUANTITIES

8.1 Are the two escape rates different?

An important question to address is whether the two escape rates $k_w, k_b$ are identical or not. To answer this question we eliminate the colour changing process as follows.

With an all white initial condition, some particles will change colour an odd number of times and will thus exit coloured black. For such a trajectory, we pretend that it began as a black particle and never changed colour. In this manner we eliminate the (colour changing) reaction in favour of changing the initial condition to a mixture of black and white particles, which then merely propagate retaining their colour labels. This procedure is implemented by taking trajectories which escape black, changing their colour histories to all black, and then storing the resulting data. The result (fig. 8.1) clearly shows that the escape rates vary with the initial distribution of particles of each colour, and hence that they are distinct from one another in a non-trivial way. It should be noted that the exponential decay starts after $t \approx 7$ (see figs. 5.2, 5.3). The white escape rate does not vary noticeably with the location of the colouring region, and remains at the uncoloured escape rate.
of $\gamma = 0.662$.

Two of the data points in fig. 8.1 are shown in full in fig. 8.2. The early time behaviour is distorted due to the procedure followed (namely that the initial condition is effectively changed to a mixture of black and white particles as described above).

8.2 Links between microscopic and macroscopic quantities

Dynamical randomness is characterized by a positive Kolmogorov-Sinai entropy $h_{KS}$. But dynamical randomness has its origins in sensitivity to initial conditions. This sensitivity is characterized by the Lyapunov exponents $\lambda_i$, which describe the rate of separation of initially infinitesimally displaced trajectories. Our system has one positive Lyapunov exponent $\lambda$. The other three are $-\lambda$, 0, and 0 since this system (for trapped trajectories) satisfies Liouville's theorem of phase space density conservation and is time-reversal symmetric.

In closed systems these quantities are related by Pesin's theorem

$$h_{KS} = \sum_{\lambda_i > 0} \lambda_i . \quad (8.2.1)$$

For open systems the formula

$$\gamma = \sum_{\lambda_i > 0} \lambda_i - h_{KS} , \quad (8.2.2)$$

where $\gamma$ is the escape rate, has been shown to hold for several classes of systems and is thought to be true quite generally ([1, 2, 6, 11, 12]). In particular it holds for our (uncoloured) system. Equation (8.2.2) can be
Fig. 8.1: Negative of the black escape rate. Each diamond is the slope of the logarithm of the number of black particles in the system as a function of time. Each diamond represents a colouring region on disk 1 centered on the $\theta$ location of the diamond and extending one diamond to either side. $\theta \in [0, \pi/6]$ only shown because of symmetry. The escape rate for the uncoloured system (fig. 1.2) is 0.662, and this is also the value for the white escape rate for all of the data shown.
8. *Escape rates & links between microscopic and macroscopic quantities* 64

![Graphs showing black density over time for different regions](image)

*Fig. 8.2:* The logarithm of the number of black particles in the system at time $t$ for colouring regions centered at $\theta = 0.035$ (left plot) and $\theta = 0.291$ (right plot) on disk 1. These correspond to the 4th and 33rd symbol from the left in fig. 8.1 respectively. Abscissa in arb. units.

interpreted as saying that dynamical randomness inhibits escape from the repeller.

The uncoloured escape rate is approximately $\gamma = 0.662$ for our system (see fig. 1.2). The uncoloured Kolmogorov-Sinai entropy is approximately $h_{KS} = \ln 2$ (see section 6.2). Thus (8.2.2) yields the positive Lyapunov exponent as $\lambda = 1.36$. This Lyapunov exponent can be (approximately) calculated directly from the paths of trapped trajectories; this was not attempted. These numerical values are for the case $R = 3.11a$.

These relations are important because they link a macroscopic quantity, namely the escape rate, with microscopic quantities, namely the Kolmogorov-Sinai entropy and the Lyapunov exponents. For the coloured case, we have a qualitative relationship between macroscopic quantitites, namely the black escape rates, and microscopic quantities, namely the coloured Kolmogorov-
Sinai entropies and the invariant density of the repeller, as shown in fig 8.3. It should be noted that the resolution of the Kolmogorov-Sinai entropy plot and of the black escape rate plot is very low compared to the repeller. Hence the fine detail of the repeller, assuming this detail is present in the other plots, is obscured.

Fig. 8.3: Fig. 8.1 (black escape rate) and the $N = 7$ curve of fig. 7.1 (coloured $h_{KS}^r$) superimposed to scale. Fig. 2.13 (the repeller density as a function of theta) is also shown, but not to scale. Abscissa label omitted.
9. COLOUR CORRELATION FUNCTION

9.1 Definition and a modification of the system

Fig. 9.1: Periodic extension of the three disk system.

As a second example of a 'reactive' Lorentz gas, we extend our three disk system to a periodic arrangement of disks on a triangular lattice (fig. 9.1). A
Fig. 9.2: Appropriate boundary conditions to revert back to a three disk domain.
A coloured region and three escaping/reinjected pairs of trajectories are shown. Each pair is coded differently (circle, square, blank) to clearly identify them.
particle introduced into the system now moves from one three disk region to another instead of escaping the system. If we colour each three disk domain in the same manner (for example, by placing colouring regions in pairs $\pi$ radians apart on every disk as in fig. 9.2) then it is sufficient for the correlation function to use one three disk domain with the appropriate boundary conditions (fig 9.2). In this manner we obtain a trajectory of arbitrary length in a single three disk domain. Note that escape has been eliminated so that the only phenomenological process left is the colour changing reaction. The phenomenological picture is

$$B = \frac{k}{k} \mathcal{W},$$

(9.1.1)

which leads to the expressions

$$\frac{dn_B}{dt} = k n_w - k n_B$$

(9.1.2)

$$\frac{dn_w}{dt} = k n_B - k n_w$$

(9.1.3)

for the number of black and white particles present in the system. Adding (9.1.2) to (9.1.3) yields

$$n_B + n_w = \text{const}$$

(9.1.4)

at all times.

The disks are separated enough to permit a class of trajectories which never strike a disk; these trajectories move in a straight line forever. When the disks are close enough together to prevent this class of trajectories, it is known that the system has strong ergodic properties [3]. However, even when the disks are far enough apart to permit this class of straight line trajectories
(which is the case for our system), this class has zero measure in the class of all trajectories, and the other trajectories sample the available phase space ergodically. This implies that there is an equilibrium colour distribution of half black and half white, each half being distributed equally throughout the system.

We are interested in the rate of return to (colour) equilibrium after the system is perturbed. Denoting the deviation from equilibrium by $\delta$, we have

$$\delta n_B = n_B - \langle n_B \rangle$$  \hspace{1cm} (9.1.5)$$

so that

$$\frac{d\delta n_B}{dt} = \frac{dn_B}{dt} = k (n_W - n_B)$$  \hspace{1cm} (9.1.6)$$

Observe that

$$n_W - n_B = = \delta n_W + \langle n_W \rangle - \delta n_B - \langle n_B \rangle$$  \hspace{1cm} (9.1.7)$$

$$= -2\delta n_B.$$  \hspace{1cm} (9.1.8)$$

Hence

$$\frac{d\delta n_B}{dt} = -2k\delta n_B$$  \hspace{1cm} (9.1.9)$$

$$\Rightarrow \delta n_B(t) = c_0 e^{-2kt}$$  \hspace{1cm} (9.1.10)$$

This immediately implies that the phenomenological description is simple exponential decay.

To probe this relaxation to equilibrium after a small perturbation, we invoke Onsager's regression hypothesis [13], which says that the relaxation of macroscopic non-equilibrium disturbances is governed by the same laws as the
decay of correlations with time due to spontaneous microscopic fluctuations in the equilibrium system. Hence we shift our attention to fluctuations in the equilibrium system, and we can use one long trajectory because of the ergodic nature of the system.

We take the colour correlation function as

\[
\frac{1}{4} c(t) = \frac{1}{N} \sum_{t' = 0}^{N} \delta\theta(t')\delta\theta(t + t') ,
\]

(9.1.11)

where \( \delta\theta(t) = \theta(t) - \langle \theta(t) \rangle_{\text{avg}} \),

(9.1.12)

where \( \theta(t) = \begin{cases} 0 & \text{if white at time } t \\ 1 & \text{if black at time } t. \end{cases} \)

(9.1.13)

Since \( \langle \theta(t) \rangle_{\text{avg}} = 1/2 \), \( \delta\theta = \pm 1/2 \) and so multiplying through by four gives

\[
c(t) = \frac{1}{N} \sum_{t' = 0}^{N} \xi(t')\xi(t + t') ,
\]

(9.1.14)

where \( \xi(t) = \begin{cases} -1 & \text{if white at time } t \\ 1 & \text{if black at time } t. \end{cases} \)

(9.1.15)

Notice that the choice of sign for the colours does not affect the correlation function. Expression (9.1.14) is to be thought of as a time average over a long trajectory. The times \( t' \) are chosen generically to provide a representative sample of the particle's behaviour. The correlation function should not exhibit a marked dependence on the particular choices of \( t' \).

### 9.2 Small colouring region behaviour

The exponential decay expected on a phenomenological level is borne out for a small colouring region (upper plots of fig. 9.3). What is the microscopic
Fig. 9.3: The upper left plot shows the correlation function for a colouring region on disk 1 spanned by $\theta \in [\pi/6 \pm 0.02]$. It also shows the correlation function obtained by sampling from the distribution (upper left plot of fig. 9.4) alone. The middle left plot shows the same curves, but for a colouring region covering all of disk 1. The upper curve is the actual correlation function, while the lower curve is the correlation function obtained by sampling from the distribution alone. The lower left plot shows the same curves again, but for a colouring region covering all three disks. The matching plots on the right hand side show the logarithms of these curves.
Fig. 9.4: The distribution of times between colour changes. The upper left plot is for a colouring region on disk 1 spanned by $\theta \in [\pi/6 \pm 0.02]$, the upper right plot is for the case where all of disk 1 is coloured, and the lower left plot is for the case where all three disks are completely coloured. The lower right plot shows the early time blowup of these distributions (drawing attention to their initial zero value). Abscissa in arb. units.
basis for this? For a small enough colouring region there are essentially no correlations between successive colour changes since the trajectory is quickly randomized. The distribution of times between successive colour changes follows the exponential (inter-arrival time) distribution of a Poisson process (the upper left plot of fig. 9.4). Given that the correlation function decays as

$$\frac{dc(t)}{dt} = -2kc(t),$$

(9.2.16)

we can ask what the decay constant $k$ is. We would expect it to be related to the stationary distribution of the system – namely to the collision frequency profile as a function of $\theta$. Indeed, the decay constant $k$ is proportional to the collision frequency profile integrated over the $\theta$ range of the coloured region. This value turns out to be independent of $\theta$ (not shown).

### 9.3 A simple model

One subtle aspect of the correlation function will now be discussed. It is observed that the distribution of times between successive colour changes remains zero for a finite interval (lower right plot of fig. 9.4). This occurs because there is a minimum time interval ($\geq 0$) between colour changes. This is easy to see for a small colouring region centered at $\theta = \pi/6$ on disk 1 since the particle must either travel to the boundary of the system to be reinjected or must collide with other disks before returning to this colouring region. That there is still a minimum time for arbitrary colouring regions is harder to see, but becomes clear after a careful examination of fig. 9.2. This minimum time has an interesting implication for $c(t)$, namely that its
9. Colour Correlation Function

early time behaviour is linear and not exponential. This shown by way of some simple models. Unfortunately this effect is difficult to detect in the correlation functions themselves since any smooth function is locally approximately linear (think of performing a Taylor expansion about a particular point). However, the linearity mentioned above is rigorous.

First a model is considered which has an interesting feature – namely the existence of straight line segments in the correlation function. In this model the trajectory can only change colour at times $nT$, where $n$ is a postive integer; the probability of it actually doing so is $p$. Hence the particle's colour remains constant within the time blocks $[nT, (n + 1)T]$, for integer $n$. This defines the model. However, if one would like a physical system to think of, we can imagine the following (fig. 9.5). A ball is thrown straight down to the lower surface and then it bounces back up and is caught at the upper surface. The ball is then positioned at random along the upper surface before being thrown again. This is repeated, with the ball changing colour only when it strikes the colour region. (The zero time value must taken when the ball first strikes the lower surface.)

To compute the correlation function for this model system, we proceed as follows. We compute just one term in the sum (9.1.14). Since (9.1.14) is an average over its component terms, it suffices to compute one generic (average) term. $t'$ is chosen uniformly in $[0, T]$ (by taking $t'$ modulo $[0, T]$ if necessary). First restrict $t \in [0, T]$ (fig. 9.6). Denote by $\mathcal{P}$ the probability of the event ($t < T - t'$). Then

$$\xi(t')\xi(t + t') = \mathcal{P} + (1 - p)(1 - \mathcal{P}) - p(1 - \mathcal{P}). \quad (9.3.17)$$
Fig. 9.5: Physical picture for the constant time block correlation model. The coloured region is shown as striped.

Fig. 9.6: Considerations to compute the correlation function for $t \in [0, T]$ in the constant time block model.
9. Colour Correlation Function

But \(T - t'\) is also uniform in \([0, T]\), so

\[
\mathcal{P} = \frac{T - t}{T}.
\]  

(9.3.18)

Hence

\[
c(t) = 1 - \left(\frac{2p}{T}\right) t \quad \text{for} \quad t \in [0, T].
\]  

(9.3.19)

---

**Fig. 9.7:** Considerations to compute the correlation function for \(t \in [(i-1)T, iT]\) in the constant time block model.

In general, restrict \(t \in [(i-1)T, iT]\) (fig. 9.7). Denote by \(\mathcal{P}\) the probability of the event \((t < iT - t')\). Then

\[
\xi(t')\xi(t + t') = \mathcal{P} \left[ \sum_{j=0}^{i} (-1)^j p^j (1 - p)^{i-j} \binom{i}{j} \right] + (1 - \mathcal{P}) \left[ \sum_{j=0}^{i+1} (-1)^j p^j (1 - p)^{i+1-j} \binom{i+1}{j} \right].
\]  

(9.3.20)

But \((iT - t')\) is uniform in \([(i-1)T, iT]\), so

\[
\mathcal{P} = \frac{iT - t}{T}.
\]  

(9.3.22)

Hence

\[
c(t) = \frac{t}{T} \sum_{j=0}^{i} (-1)^j p^j (1 - p)^{i-j} \binom{i}{j}
\]  

(9.3.23)

\[
+ 1 - i + \frac{t}{T} \sum_{j=0}^{i+1} (-1)^j p^j (1 - p)^{i+1-j} \binom{i+1}{j}
\]

(9.3.24)

for \(t \in [(i-1)T, iT]\).
Fig. 9.8: Physical picture for the variable time block correlation model. The coloured region is shown as striped.

Do these linear segments persist when we add variability to the model? The physical picture (fig. 9.8) changes so that the ball is caught and rethrown somewhere in the fuzzy region. We now let each time block vary in length from $T_{\text{min}}$ to $T_{\text{max}}$. We denote this length by $T$, where $T$ is now no longer a constant. $t'$ is still uniform over $[0, T]$, and restricting $t \in [0, T_{\text{min}}]$ gives (fig. 9.9)

$$\xi(t')\xi(t + t') = 1 - \left(\frac{2p}{T}\right)t$$

(9.3.25)

again. Now, though, to get $c(t)$ we must average over $T$, so that

$$c(t) = 1 - 2pt \left\langle \frac{1}{T} \right\rangle_{\text{avg}} \quad \text{for } t \in [0, T_{\text{min}}].$$

(9.3.26)

We see that the initial linearity remains.
9. Colour Correlation Function

Fig. 9.9: Considerations to compute the correlation function for \( t \in [0, T_{\text{min}}] \) in the variable time block model.

9.4 Large colouring region behaviour

The decay of the correlation function for large colouring regions is not exponential (middle and lower plots of fig. 9.3). Hence the simple rate law given by the phenomenological picture is not valid. There are several possible reasons for this breakdown, and it is worth studying their effects. We will look at two effects, namely deviation from a Poisson process and correlation of neighbouring time intervals.

As the particle moves in the system it generates a coloured time history, which at the simplest level consists of the times at which the particle changes colour. This is all the information needed to compute the correlation function. We now focus on the time intervals between these colour changing events (upper right and lower left plots of fig. 9.4).

These are the inter-arrival times in the language of Poisson processes, and for a Poisson process have an exponential distribution. The inter-arrival time distribution can be collected along with the correlation function, and our first interest is in whether anything is lost by only having this distribution. By only having the distribution one loses knowledge about the arrangement of the time intervals in relation to one another. The same distribution can give
9. Colour Correlation Function

rise to very different correlation functions depending upon this arrangement. To test whether or not this arrangement is important for the system of interest, one can compare the correlation function for the system (with the full dynamics) to the correlation function obtained by sampling from the distribution alone. The results show that the arrangement of the time intervals for the full system are generic and have no effect on the correlation function (fig. 9.3) (except in the middle plots), so that any deviation from exponential decay must come from the actual form of the inter-arrival time distribution, not its time ordering.

Now we look at the shape of the inter-arrival time distribution for various colouring regions (fig. 9.4). We observe that the distribution has structure at short times, and in the small colouring region case (the upper left plot in fig. 9.4) becomes smoothly exponential at long times. This is reminiscent of the initial events section (chapter 5). For large colouring regions trajectories are likely to change colour often, either from one reinjection to the next or within the same reinjection. These multiple collisions with colouring regions are too frequent for the trajectory to be randomized — there are correlations between successive colour changes which introduce structure into the distribution of times between colour changes. This structure is extremely pronounced in the lower left plot of fig. 9.4, where a colour change and a collision event with a disk are synonymous. This structure and the severe breakdown of the phenomenological picture is the result of two time scales being nearly equal — namely the time scale for colour change and the time scale for randomization. It is the relative size of these two time scales which determines the validity of the assumptions implicit in a Poisson process description and hence in a
phenomenological picture.
10. CONCLUSION

The systematic use of N-cylinders allowed for a detailed understanding of the Lorentz gas system under study. This generating partition was particularly well suited for studying the Kolmogorov-Sinai entropy and its finite and coloured varieties. The N-cylinders are approximations to the repeller, which underlies the chaotic nature of the system. The repeller and the coloured Kolmogorov-Sinai entropies were shown to be qualitatively similar to macroscopic quantities, namely the black escape rate for various colouring regions. This colouring introduced a label by which to observe particular classes of trajectories. These classes were observed to be substantially different from one another. Phenomenological pictures were introduced to give physical insight. The breakdown of these pictures, namely at early times for the three disk system and for large colouring regions in the periodic system, was examined. Some models which, in both the Kolmogorov-Sinai and correlation function\(^1\) sections, replaced the exact dynamics with statistical schemes were introduced to gain insight and to draw attention to particular aspects of the systems.

Increasing the value of \(R\) attenuates the N-cylinders. For example, the lower left plot of fig. 2.10 would have three narrow peaks with larger valleys in

\(^1\) including the 'model' where we sample from the distribution alone
Correspondingly, the peaks in figs. 7.1 and 8.1 would be compressed and the valleys larger. The exponential decay constants would increase as well. The general features remain unchanged.
A. USE OF PROBABILITY

The system of interest has two degrees of freedom and is deterministic. Hence one might question the need for a probabilistic description at all. There are two compelling reasons to use probability. The first is a statistical mechanics reason while the second is a nonlinear dynamics reason.

The first has to do with the type of information we are interested in obtaining from the system. Our goal is to find adequate global features to characterize the behaviour of the system. Characteristics of individual trajectories, such as their lifetime, fluctuate strongly with the initial condition. For a statistical description we must consider an ensemble of initial value; in other words an initial distribution of particles.

The second reason is due to the sensitive dependence of the dynamics on the initial conditions. Two generic trajectories starting from infinitesimally separated initial conditions will deviate from each other exponentially quickly and before long lose their resemblance. This means that if the initial condition is not known exactly then the resulting behaviour becomes unpredictable after a short time. In general it is not possible to follow a trajectory analytically, and hence the only recourse is to a statistical description.
B. ACCESSIBILITY CONDITION

The system changes in nature at $R = 3a$. For $R > 3a$ any point on one of the disks is accessible from any point on either of the other two disks by a direct path, provided all points are within their $\theta$ bounds as given above (1.4.6).

This accessibility condition is violated in fig. B.1 for the dashed line; it must pass through disk 1 to reach disk 2. The closest allowed path to this is the tangent path, shown as a solid line. The dashed line case is easily seen to be the most stringent, so that if it is allowed then all paths become possible. The smallest $R$ value required to satisfy this condition is calculated as follows.

A point on disk 1 is $(a \cos \theta_1, a \sin \theta_1)$. A point on disk 2 is $(R + a \cos \theta_2, a \sin \theta_2)$. The equation of the line joining them is

$$\frac{y - a \sin \theta_1}{x - a \cos \theta_1} = \frac{a \sin \theta_2 - a \sin \theta_1}{R + a \cos \theta_2 - a \cos \theta_1}. \quad (B.0.1)$$

Disk 1 is given by the equation $x^2 + y^2 = a^2$. Let us differentiate implicitly.

$$2x + 2y \frac{dy}{dx} = 0 \quad (B.0.2)$$

$$\Rightarrow \frac{dy}{dx} = \frac{-x}{y} = -\frac{a \cos \theta_1}{a \sin \theta_1} = -\cot \theta_1 \quad (B.0.3)$$

Now we equate the slope of the tangent (RHS of (B.0.1)) with the slope of
Fig. B.1: Violation of accessibility condition
the line (RHS of (B.0.3)) and set \( \theta_1 = \pi/3, \theta_2 = \pi \), giving

\[
\frac{-a\sqrt{3}/2}{R - a - a/2} = -\frac{\sqrt{3}}{3}
\]

\[\Rightarrow R = 3a . \quad (\text{B.0.4})\]
C. GENERALIZATION FOR ARBITRARY R

The above calculation yields the correct result for a 3 disk system with the disks far enough apart. However, it is desirable to work through a more general case, which is presented below. Instead of limiting ourselves to the range (1.4.6), we allow $\theta$ to be arbitrary, in which case our focus changes from hitting disks within the range (1.4.6) to merely hitting them at all. Thus the critical trajectories are the tangents.

Starting from the point $(a \cos \theta, a \sin \theta)$ on disk 1, an arbitrary straight line is given by

$$\frac{y - a \sin \theta}{x - a \cos \theta} = \gamma .$$  \hspace{1cm} (C.0.1)

Let us find the points in common between this and disk 2, which is given by

$$y^2 + (x - R)^2 = a^2.$$  

The common solution is given by eliminating $y$, yielding

$$x^2 \left[ \gamma^2 + 1 \right] + x \left[ 2\gamma a \sin \theta - 2\gamma^2 a \cos \theta - 2R \right] +$$

$$\left[ a^2 \sin^2 \theta + \gamma^2 a^2 \cos^2 \theta - 2\gamma a^2 \cos \theta \sin \theta + R^2 - a^2 \right] = 0 .$$  \hspace{1cm} (C.0.2)

This is a quadratic in $x$, whose discriminate is

$$\gamma^2 \left[ 8Ra \cos \theta - 4a^2 \cos^2 \theta - 4R^2 - 4a^2 \right]$$

$$+ \gamma \left[ 8a^2 \sin \theta \cos \theta - 8Ra \sin \theta \right] + \left[ 4a^2 \cos^2 \theta \right] = 0 .$$  \hspace{1cm} (C.0.3)
For the tangential solution, we need the solution to collapse down to one root, hence the discriminate must be set equal to zero. The discriminate, as written, is a quadratic in $\gamma$, whose solution is

$$\frac{R \sin \theta - a \sin \theta \cos \theta \pm \sqrt{R \sqrt{R^2 - 2a \cos \theta}}}{a(\cos \theta - R/a + 1)(\cos \theta - R/a - 1)} = \gamma \pm . \tag{C.0.4}$$

The two values are the upper and lower paths as shown in fig. C.1. We must restrict the line of slope $\gamma$ to fall outside of disk 1, so that we must watch out for the critical value $\gamma = -\cot \theta$, which is the tangent condition for disk 1.

These expressions are in terms of $\theta$ and $\gamma$. We must transform into the $\theta-\phi$ coordinates of before. This is straightforward, giving

$$\phi = \pi - \theta - \arctan \gamma . \tag{C.0.5}$$

The critical value $\gamma = -\cot \theta$ becomes, of course, $\phi = \pi \pm \pi/2$. 

---

**Fig. C.1**: Generalized extrema of cell -12
The other curves can be obtained from the first by symmetry arguments similar to those previously considered. For example, fig. C.2 shows that

\[ \phi_{-12}^{\text{out}}(2\pi - \theta) = 2\pi - \phi_{-12}^{\text{in}}(\pi + \theta) \]  
(C.0.6)

\[ \Rightarrow \phi_{-12}^{\text{out}}(\theta) = \phi_{-12}^{\text{out}}(\pi - \theta). \]
D. COMMENTS ON NUMERICAL WORK

Some comments are in order whenever one studies chaotic systems numerically. Stephen Smale [14], for instance, profoundly questions this practice, which is so common today in the sciences.

It should be noted first that our system has no attractor. Hence it is fundamentally different from most chaotic systems in the sense that an arbitrary initial condition will leave the system quickly.

The systematic use of N-cylinders, which are also derived analytically as shown above, allows us to refine information in a consistent manner.

All results were obtained using double precision reals. Generic examples were compared with their single and quadruple precision counterparts. The differences are vanishingly small. Furthermore, all the plots shown have negligible error bars in the sense that using more particles (trajectories) does not change the curves on the scales shown. However, it is important to realise that in determining the slope of, say, fig. 1.2, the very long time section is noisy due to the small number of trajectories remaining in the system at these times. Hence this region is not of much value in calculating the slope. This implies that the data shown in fig. 8.1 is restricted to approximately $N \leq 20$ in the N-cylinder picture. These implications are discussed in section 8.2.
While it is certainly true that any particular trajectory deviates from its exact path due to round-off error, all of the results are averaged over an ensemble of initial conditions, so that a 'δ shadowing theorem' applies.
E. ADDITIONAL DETAILS OF CHAPTER 2.1

E.1 Comparison with [1]

This is to be compared with [1, (3.1)], which is not at first glance the same. To show that it is indeed identical to (2.1.3) we manipulate (2.1.3) as follows.

\[
\phi = \pi + \theta - \arctan \left( \frac{\sin \theta}{\cos \theta - \frac{R}{a} + 1} \right) \\
= \pi + \arctan(\tan \theta) - \arctan \left( \frac{\sin \theta}{\cos \theta - \frac{R}{a} + 1} \right) \\
[15] = \pi + \arctan \left( \frac{\tan \theta - \frac{\sin \theta}{\cos \theta - \frac{R}{a} + 1}}{1 + \frac{\sin \theta \tan \theta}{\cos \theta - \frac{R}{a} + 1}} \right) \\
= \pi + \arctan \left( \frac{\frac{\tan \theta - \frac{R}{a} \tan \theta}{\cos \theta - \frac{R}{a} + 1}}{\frac{\sin \theta \tan \theta + \cos \theta - \frac{R}{a} + 1}{\cos \theta - \frac{R}{a} + 1}} \right) \\
= \pi + \arctan \left( \frac{\sin \theta - \frac{R}{a} \sin \theta}{\sin^2 \theta + \cos^2 \theta - \frac{R}{a} \cos \theta + \cos \theta} \right) \\
= \pi + \arctan \left( \frac{\sin \theta(1 - \frac{R}{a})}{1 + \cos \theta(1 - \frac{R}{a})} \right) \\
= \pi + \arctan \left( \frac{\sin \theta}{\cos \theta - \frac{a}{R-a}} \right). \\
\]

(E.1.1)

93
Fig. E.1: Geometric considerations for $\phi_{1/2}$
E.2 Explicit derivation of $\phi_{\pm}^{1-2}$ and $\phi_{\pm}^{1-2}$

The equation of the line from the point $(a, 0)$ on disk 1 to the generic point $(R + a \cos \theta, a \sin \theta)$ on disk 2 is (fig. E.1)

$$\frac{y}{x - a} = \frac{a \sin \theta}{R + a \cos \theta - a} , \quad (E.2.2)$$

making the angle $\chi$ satisfy

$$\tan \chi = \frac{a \sin \theta}{R + a \cos \theta - a} . \quad (E.2.3)$$

The angle we want is (fig. E.1)

$$\phi^{out} = 2\pi - \theta + \chi$$
$$= 2\pi - \theta + \arctan \left(\frac{\sin \theta}{R/a + \cos \theta - 1}\right) . \quad (E.2.4)$$

This should give us the same result as (2.1.3), which was derived using symmetry. Let us verify this, starting with (2.1.3).

$$\phi_{-}^{1-2}(\pi - \theta) = \pi + (\pi - \theta) - \arctan \left(\frac{\sin(\pi - \theta)}{\cos(\pi - \theta) - R/a + 1}\right)$$
$$= 2\pi - \theta - \arctan \left(\frac{\sin \theta}{- \cos \theta - R/a + 1}\right)$$
$$= 2\pi - \theta + \arctan \left(\frac{\sin \theta}{R/a + \cos \theta - 1}\right)$$
$$= \phi_{-}^{1-2}(\theta) . \quad (E.2.5)$$

The equation of the line from the point $(a \cos(\pi/3), a \sin(\pi/3))$ on disk 1 to the generic point $(R + a \cos \theta, a \sin \theta)$ on disk 2 is (fig. E.2)

$$\frac{y - \sqrt{3}a/2}{x - a/2} = \frac{a \sin \theta - \sqrt{3}a/2}{R + a \cos \theta - a/2} . \quad (E.2.6)$$
Fig. E.2: Geometric considerations for $\phi^{1/2}$
making the angle $\chi$ satisfy

$$\tan \chi = \frac{-a \sin \theta + \sqrt{3}a/2}{R + a \cos \theta - a/2}.$$  \hfill (E.2.7)

The angle we want is (fig. E.2)

$$\phi^{out} = 2\pi - \theta - \chi = 2\pi - \theta - \arctan \left( \frac{-\sin \theta + \sqrt{3}/2}{R/a + \cos \theta - 1/2} \right).$$  \hfill (E.2.8)
BIBLIOGRAPHY


[14] *Great Problems*, a lecture given as part of the Arnol'd Conference at the Fields Institute, University of Toronto, June 19, 1997.


