Deterministic Volume Approximation

by

Rares Cristian

and advised by

Santosh Vempala

Contents

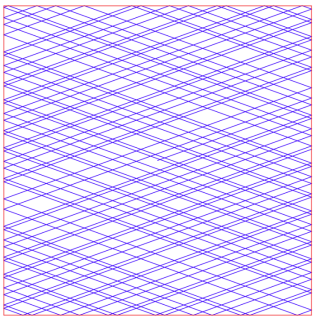
Introduction

Our research concerns computing the volume of high-dimensional polytopes. Currently, there are no known deterministic algorithms that run polynomial in the dimension, although surprisingly, there do exist efficient randomized algorithms to approximate volume. Essentially, the problem has been reduced to sampling uniformly at random from the interior of the body, with current algorithms performing this via Monte Carlo Markov Chain methods. See [1] and [2] for details on previous approaches.

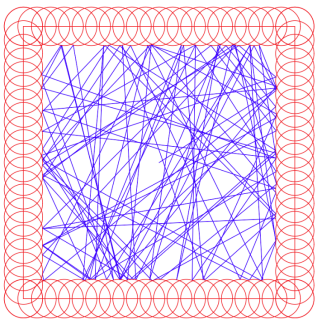
We investigate whether the notion of chaos can serve as a substitute for randomness in this setting. There is a distinction to be made here between chaos and randomness. Given the initial state of a system, we cannot predict the future state of a random process. On the other hand, a chaotic one is fully deterministic. But this process is sensitive to initial conditions: any small change in initial conditions will result in vastly different orbits.

In particular, we consider the orbit created by the free motion of a point particle inside the polytope with mirror-like reflections off the boundary. Figure (a) depicts the trajectory of a point particle in a square. However this system is not chaotic as translating the initial point will result in largely the same result. To remedy this, we introduce an inward curvature to the boundary. Figure (b) depicts the trajectory of the same initial point, only with circles placed around the boundary.

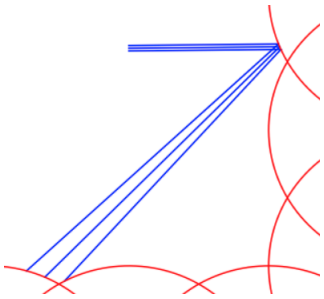
(a) Flat Boundary



(b) Dispersing Boundary



Intuitively, this curvature generates a dispersing effect, causing particles that were initially near each other to spread out as time goes on. Hence, we will alter the initial polytope and compute this reflection sequence in the curved version.



We are interested in proving that for any subset A of the polytope, the trajectory of a particle will spend time in A proportional to its volume. In particular, this is described by the sequence

$$S_n = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_A(x_i)$$

which counts the fraction of points within the first n reflections which lie in the set A. Ergodicity of the system would imply that in the limit this sequence converges appropriately. The classical theory does not derive quantitative bounds on convergence. This paper show that this sequence indeed converges at an exponential rate. To do so, we will analyze the billiard system and the proofs of ergodicity and of exponential decay of correlations for dispersing billiards. We will first focus our attention on the case of the simplex, rather than general polytopes.

Additionally, there's a potential issue here in implementation. At each iteration of this algorithm, we lose a small amount of precision due to floating point errors. In a chaotic system, these small errors can lead to a vastly different trajectory than the original. To remedy this, we will investigate the notion of *shadowing*. Every pseudo-trajectory, which can be thought of as a numerically-computed trajectory with rounding errors introduced at each step, stays uniformly close to some true trajectory. That is, the pseudo-trajectory we generate is *shadowed* by a true one.

Ball Reflections

Here, we provide an algorithm for efficiently computing the intersection between the trajectory of the particle and the boundary of the billiard table in the dispersing version.

2.1 The Problem

We are given a hyperplane H containing point q, and perpendicular to vector a. That is,

$$H = \{x \in \mathbb{R}^n : a^T(x - q) = 0\}$$

For simplicity, we will assume that q is the origin. A simple translation will bring H into this position. Let $B = \{b_1, \ldots, b_{n-1}\}$ be an orthogonal basis whose span is H. We set a grid on top of H, such that each grid point can be described as

$$(\delta \cdot \lambda_1 b_1, \ldots, \delta \cdot \lambda_{n-1} b_{n-1})$$

for some fixed $\delta \in \mathbb{R}$ which we call the grid width, and $\lambda_i \in \mathbb{Z}$. At each gridpoint, we place a ball of radius R. Given a ray defined by a starting point p and a direction vector d, we wish to find the smallest positive value of t such that $p + dt$ intersects some sphere.

We must choose the radius R so that any line not parallel to H intersects at least one ball. Thus, the spheres must completely cover the hyperplane. Since H is tiled by $(n-1)$-dimensional cubes of side length δ, it suffices to choose R so that each point in such a cube is at a distance at most R from one of its corners (which is where the balls will be placed). The farthest point from any corner is the very center of the cube, so the radius must be at least $\frac{\delta}{2}\sqrt{n-1}$.

2.2 Broad Overview

For a hyperplane H, define function f_H where $f_H(x)$ is the gridpoint on H closest to x. We restrict our attention to f_H along the ray $l(t) = p + dt$. We can divide l into maximal contiguous segments $[a, b]$ for which $f_H(x)$ remains constant (say with value g) for all $x \in$

$[a, b]$ and we say that g corresponds to segment $[a, b]$. Let b be the center of the first ball intersecting l. Clearly, b corresponds to some segment. We will iterate through the segments until we find b.

2.3 Preliminaries

First, we must build the grid. We will not explicitly create the points on the grid, but rather we will see that we can generate the gridpoints we care about on the fly. However, we do need to have the basis $\{b_1, \ldots, b_{n-1}\}$ on hand. We may choose any basis $\{a, v_1, \ldots, v_{n-1}\}$ for $\mathbb{R}^n$ and afterwards apply Gram-Schmidt orthogonalization which will keep the first entry a in the basis. Then, the remaining vectors will all be perpendicular to a and will thus span H.

Next, we show how to compute $f_H(x)$. Let x_H be the projection of x onto H. The distance from any gridpoint $y \in H$ is $\sqrt{\|x - x_H\|^2 + \|y - x_H\|^2}$, so it suffices to find y minimizing distance to x_H since $\|x - x_H\|^2$ is independent of y. To do so, we may look at x_H represented in the coordinates of the grid. That is, $B^{-1}x_H$, where $B = \begin{bmatrix} b_1 & b_2 & \ldots & b_{n-1} & a \end{bmatrix}$ (note that the last entry is guaranteed to be 0, since x_H is on the hyperplane, and so has no perpendicular component). The coordinate point closest to $B^{-1}x_H$, say y^*, can be found by rounding each entry to the nearest multiple of δ. Then, $f_H(x) = By^*$. This process takes $O(n^2)$ time. Also, note that $B^{-1} = B^T$ since B is orthogonal by construction.

2.4 Finding Next Segment

In this section, we focus on finding the next segment along a ray $l(t) = p + dt$ given some starting segment. Without loss of generality we may assume that $t = 0$. We would like to find the smallest $t > 0$ such that $f_H(l(t))$ is different from $f_H(l(0))$. We will see that it is sufficient to find this gridpoint, $f_H(l(t))$ in the coordinates of the grid, and that there is no need to convert it to standard coordinates. So, we may assume that we are given the ray in this form.

2.4.1 Binary Search

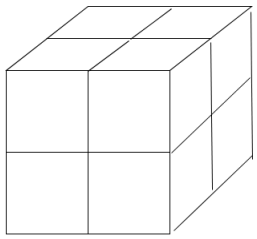
We may easily solve this problem via binary search. Say we are considering whether the answer is less than t, or greater than t. At each iteration, we may compute the nearest gridpoint of $p + dt$, which can be done in linear time by rounding each coordinate to the nearest multiple of δ. If the gridpoint closest to p is different from the one closest to $p + dt$, we may restrict our attention to values smaller than t, and to values greater than t otherwise.

But on what range do we perform the search? and since this is over a continuous interval, to what precision? We are interested in the minimum distance we must travel along some

arbitrary line for the nearest gridpoint to be guaranteed to change. A trivial upper bound of this is simply the length of the diagonal of one of the cubes, $\delta\sqrt{n-1}$. Additionally, we require precision of $\delta/2$ since we only need to be able to distinguish between gridpoints which are all a distance of at least δ from each other. Thus, the binary search itself, if we disregard the initial transformation of the ray into the grid's coordinates, takes $O(n \log n)$ time to find a single segment.

2.4.2 Linear Time Approach

For the moment, let's restrict our attention to a single $(n-1)$-dimensional cube of the grid. The set of points within a cube which are closets to a particular corner is itself a cube with side length $\delta/2$ (which we will call subcubes) as shown below:

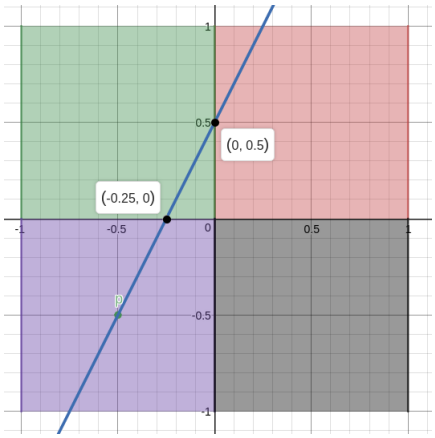


A segment is a maximal set of points on the line which lie within the same subcube. We wish to find the smallest $t > 0$ for which the ray crosses from one subcube to another. This occurs when one of the coordinates increases (or decreases) to the next multiple of $\delta/2$. There are only n such candidates (one for each coordinate), each of which we can determine as follows:

The nearest multiple of $\delta/2$ to coordinate x_i is $c_i = \lfloor \frac{x_i}{\delta/2} \rfloor (\delta/2)$. However, at $t = 0$, we may have already passed this point, and need to go to the next multiple. So, the time t_i at which the i^{th} coordinate reaches the next multiple of $\delta/2$ is

$$t_i = \begin{cases} \frac{c_i - x_i}{d_i} & \text{sign}(c_i - x_i) = \text{sign}(d_i) \\ \frac{\delta/2 - (c_i - x_i)}{d_i} & \text{otherwise} \end{cases}$$

We may simply choose the smallest t_i over all coordinates to find the starting point of the next segment. It appears that we may take a smarter approach here to compute the next segment since the current n candidates will remain to be candidates for finding the next segment. However, we will see that we need to explicitly compute $f_H(l(t))$ which takes $O(n)$ time, so there will be no significant saving in computing t_i more efficiently than in linear time. Below is an image depicting the two candidates as black points, and p in green, the starting point.



2.5 Putting it Together

Essentially, we only need to repeatedly apply the above subroutine of finding the next segment. However, there are a few things we must be careful of

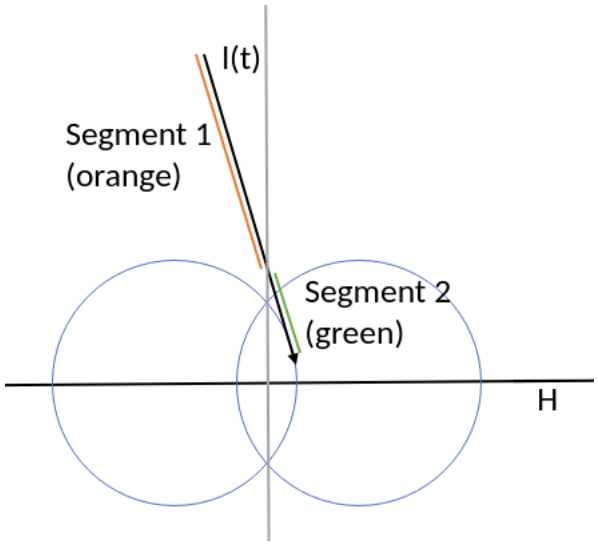
1. If the distance between $l(0)$ and H is greater than R, we must check if the ray actually intersects H. If not, we're done. Note that if the distance were less than R and the ray does not intersect H then it is still possible for l to intersect some ball. We may assume the ray is not perfectly parallel to H. So, $l(t)$ must therefore increase in distance to H as t increases. We may declare no intersection occurs if we find none by the time that the distance is greater than R.

2. If the starting point is at a distance greater than R from H, and the ray indeed intersects H, then we can skip directly to the point along l which lies at a distance of exactly R from H. This can easily be done by computing the intersection between l and the hyperplane H translated by R units in the direction a.

3. Is the starting point itself on a ball? If so, we need to be careful not to return this as our answer. Instead, we automatically move on to the next segment. The safest way to determine if we start on a ball is to keep track of the last ball we hit, or more easily, the facet this ball lies on. If we're checking for intersection with this same facet, then we are starting on a ball (this is specific to the overarching reflection algorithm).

4. Clearly, for each gridpoint y that we find, we must check if the ray intersects the ball centered at y. However, care must be taken here since we are actually computing ball-line intersection. We must additionally ensure that the corresponding value of t is positive.

5. The first ball we find which intersects l is not necessarily the answer. Consider the following example. $f_H(l(0))$ is the ball on the left, which l indeed intersects. However, that is not the first intersection with a ball.



If l intersects the ball at some point which itself is not contained within any other ball, then we have found the answer. This is true since this implies the ball intersects l somewhere on this segment, and so the answer cannot be some later segment, and we have already checked it is not any previous segment.

It is of note that each time we find a new segment, we must check if the line intersects the corresponding ball. Naively, we might convert the center of the ball into standard coordinates, requiring $O(n^2)$ time per segment for an overall runtime of $O(Sn^2)$ where S is the number of segments needed to explore. However, we may instead originally convert the line into the coordinates of the grid (plus the perpendicular component) and check for intersection within there. Intersection in the original coordinates occurs if and only if intersection occurs within the changed coordinates.

Distance between points $x, y \in \mathbb{R}^n$ is preserved when mapping them into the coordinates of the grid. Note $\|B^{-1}x - B^{-1}q\|^2 = (B^{-1}(x-y))^T(B^{-1}(x-y)) = (x-y)^T(B^{-1})^T B^{-1}(x-y) = \|x-y\|^2$ since B is unitary. Additionally, a ray intersects a ball centered at x if and only if $\min_{t\geq 0} \|x - l(t)\| \leq R$. Since distance is preserved between the two coordinate systems, this is true in the original coordinates if and only if it is true in the coordinates of the grid.

The entire algorithm may be described as follows:

1. Convert l into l' which is in the coordinates of the grid.
2. Project l' onto the hyperplane H. Let this be l''.
3. Find the next segment as described in section 2.3.2. Iteratively search for the next segment until l'' intersec‘ts the corresponding ball.
4. Convert the ball found into standard coordinates, and compute intersection with l.

Steps 1 and 4 both require $O(n^2)$ time as they perform matrix-vector multiplications. Step 3 as described in the previous section requires $O(n)$ time to find the following segment. Overall, the algorithm takes $O(n^2 + S \cdot n)$ time.

To find intersection within the polytope, we first compute the point of intersection of this ray with the polytope as if there were no balls present. We now have a segment, $[p, q]$. It is sufficient to check for intersection with balls which lie on a hyperplane H if the minimum distance between H and $[p, q]$ is at most R. So, we may potentially need to compute intersections with very few facets, unless the ray comes close to a low-dimensional flat of the polytope (such as a vertex at worst case).

Literature Review

3.1 Preliminaries

We first provide an introduction to the current approaches to approximating volume of high-dimensional polytopes $K \subset \mathbb{R}^n$, and afterwards present a connection to chaotic billiards.

3.1.1 Volume

We define the volume of K as the Lebesgue measure over K. This is a standard way of assigning measure to subsets of $\mathbb{R}^n$ and follows the standard definitions of area and volume in 2 and 3 dimensions that we are used to. We denote the volume of K as $\text{vol}(K)$.

3.1.2 Model of Computation

A key aspect of the problem lies in the representation of the set. Perhaps in the most general sense, a description of the body is not directly given, but rather access is provided by an oracle. A *membership oracle* may be given an input $x \in \mathbb{R}^n$ and returns whether or not x is contained in K. Alternatively, a *separation* oracle determines if x lies in K, and if not, provides a hyperplane separating x and K. A well-guaranteed oracle additionally provides an initial point, x_0 guaranteed to be in K, as well as bounds on the size of K: a guarantee that a ball of radius r is contained completely within K, and a guarantee that a ball of radius R fully contains K. That is, $x_0 + rB^n \subset K \subset x_0 + RB^n$.

On the other hand, we may be given an explicit description of the set. The most common example is that of a polytope. Even here, there are various ways to represent it, for example, as the convex hull of a set of points, or as the intersection of halfspaces. The latter is often more expressive, since with only a polynomial number of halfspaces, we potentially require an exponential number of vertices to describe the same polytope. Take the simple example of a cube in n dimensions. It is the intersection $2n$ halfspaces, but has 2^n corners.

3.2 Algorithms

Given only a well-guaranteed separation oracle, no deterministic polynomial-time algorithm can approximate the volume to within a factor exponential in the dimension [3]. For any sequence of n^a points there exist two different convex sets with volume ratio

$$\left(\frac{cn}{a \log n}\right)^n$$

whose oracles produce the same result to each query (for some universal constant c). Therefore, given no further information, there is no way to differentiate between the two bodies. Surprisingly, there do exist efficient randomized algorithms to compute volume given only an oracle. Volume computation is one of a few problems where randomization provably helps.

3.2.1 Overview

One major technique is to construct a series of bodies $K = K_0 \supset K_1 \supset \cdots \supset K_m$ in such a way that the volume of K_m can be easily calculated, and the ratio of volumes of K_i to K_{i+1} can be well approximated. This last step may be done by sampling uniformly at random from the interior of K_i and taking the ratio of points lying in K_{i+1}. The current state-of-the-art methods perform this sampling via random walks inside the body. The volume can be directly found as

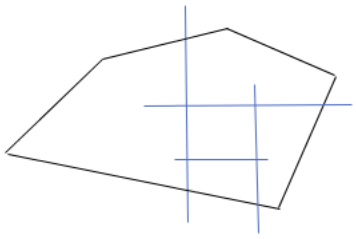
$$\text{vol}(K) = \text{vol}(K_m) \prod_{i=0}^{m-1} \frac{\text{vol}(K_i)}{\text{vol}(K_{i+1})}$$

At this point, one wishes for two things: to reduce the number of bodies K_i needed, and to reduce the number of samples required at each step. Note that the errors in approximating volume ratios are multiplicative; having fewer intermediary bodies allows us to have larger errors in approximating volume ratios, which in turn allows us to sample fewer points. At the same time, we want the ratios to be fairly large, so K_{i+1} needs to have a constant fraction of the volume of K_i. This is so we only require a polynomial number of samples to accurately approximate the ratio.

3.2.2 Constructing K_i

In addition to a convex body $K \subset \mathbb{R}^n$, say we are given that a cube of side length $r \in R$ contained in K, as well as an oracle for computing the centroid. Let e_j be an axis vector so that the width of K_i along e_j is at least r. Let z_i be the center of mass of K_i. Consider the hyperplane through z_i perpendicular to e_j. This hyperplane divides the polytope into two halves, and we let K_{i+1} be the half which contains z_0, the center of mass of K.

Figure 3.1: Centroid Cutting Plane Method



Since we choose these hyperplanes perpendicular to the axes, we terminate this process when K_m is a box, whose volume is simply the product of its side lengths. This is guaranteed to terminate once the width along each direction is at most r. It remains to bound m. But first, we require some more geometric results.

3.2.3 Geometry

Brunn-Minkowski Inequality

We begin with a general question: how does volume of a convex body change as we move along some direction θ? For simplicity, we will assume that θ is the x_1 direction. Let $K(t)$ be the $(n-1)$-dimensional "slices" of K perpendicular to θ. That is,$K(t) = K \cap \{x \in \mathbb{R}^n : x_1 = t\}$. What can we say about $K(t)$ as we vary t? We begin with a useful inequality regarding volumes of arbitrary sets, not only convex bodies.

Theorem 3.2.1 (Brunn-Minkowski Inequality)**.** *Let $A, B \subset \mathbb{R}^n$ be compact measurable sets. Then,*

$$\text{vol}(A+B)^{1/n} \geq \text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}$$

where $A + B$ is the Minkowski sum of two sets, $\{a + b : a \in A, b \in B\}$

Since $\text{vol}(\lambda A) = \lambda^n \text{vol}(A)$, we have the following equivalent version of the inequality, which we use in the next lemma.

$$\begin{aligned}
\text{vol}(\lambda A + (1-\lambda)B)^{1/n} &\geq \text{vol}(\lambda A)^{1/n} + \text{vol}((1-\lambda)B)^{1/n}, \\
&\geq (\lambda^n \text{vol}(A))^{1/n} + ((1-\lambda)^n \text{vol}(B))^{1/n} \\
&\geq \lambda \text{vol}(A)^{1/n} + (1-\lambda)\text{vol}(B)^{1/n}
\end{aligned}$$

Moreover, this implies that the volume function is $1/n$-concave with respect to Minkowski sum.

Lemma 3.2.2. *The function $(\text{vol}_{n-1} K(t))^{1/(n-1)}$ is concave.*

Proof. Let $x, y \in \mathbb{R}, \lambda \in [0,1]$ and $c = \lambda x + (1-\lambda)y$. Let $a \in K(x), b \in K(y)$. By convexity, points of the form $\lambda a + (1-\lambda)b$ all lie in K. Since they also have their first coordinate equal to the first coordinate of c, we find that $\lambda K(x) + (1-\lambda)K(y) \subset K(c)$. By the Brunn-Minkowski inequality (3.2.1),

$$\text{vol}(K(c))^{1/(n-1)} \geq \lambda \text{vol}(K(x))^{1/(n-1)} + (1-\lambda)\text{vol}(K(y))^{1/(n-1)}$$

Note that the above is in $(n-1)$ dimensions, since $K(t)$ is an $(n-1)$-dimensional slice of K.

□

Hyperplanes through the Centroid

Definition 3.2.3 (Centroid)**.** *The center of mass, or centroid, of a body $K \subset \mathbb{R}^n$ is*

$$\frac{1}{\text{vol}(K)} \int_K x dx$$

Let's take a look at a specific example: the cone. Clearly, the centroid will be on the perpendicular drawn from the apex to the base, so it remains to find the height from the base. We can integrate along the height of the cone, C, and each cross-section is an $(n-1)$-dimensional ball of radius tR/h, where R is the radius of the base, and h is the height of the cone. We say the volume of an ball of radius r in n dimensions is $f(n) \cdot r^n$, for some function f. So, the centroid is at height h^* where

$$\begin{aligned} h^* &= \frac{1}{\text{vol}(C)} \int_{t=0}^{h} t \cdot f(n-1) \left(\frac{tR}{h}\right)^{n-1} dt \\ &= \frac{f(n-1)}{\text{vol}(C)} \frac{R^{n-1}}{h^{n-1}} \int_{t=0}^{h} t^n dt \\ &= \frac{f(n-1)}{\text{vol}(C)} \frac{R^{n-1}h^2}{n+1} \end{aligned}$$

Since the volume of C is

$$\text{vol}(C) = \frac{f(n-1)R^{n-1}h}{n}$$

we find that $h^* = \frac{n}{n+1}h$. Equivalently, the centroid is at a distance of $\frac{1}{n+1}h$ from the base.

Theorem 3.2.4 (Grunbaum's Inequality)**.** *Given a convex body $K \subset \mathbb{R}^n$, any halfspace H which contains the center of gravity of K, also contains $1/e$ of the volume of K.*

We have a related result regarding the width. Define the support function of K as $h_K(x) = \max_{y \in K} \langle x, y \rangle$ for $x \in K$.

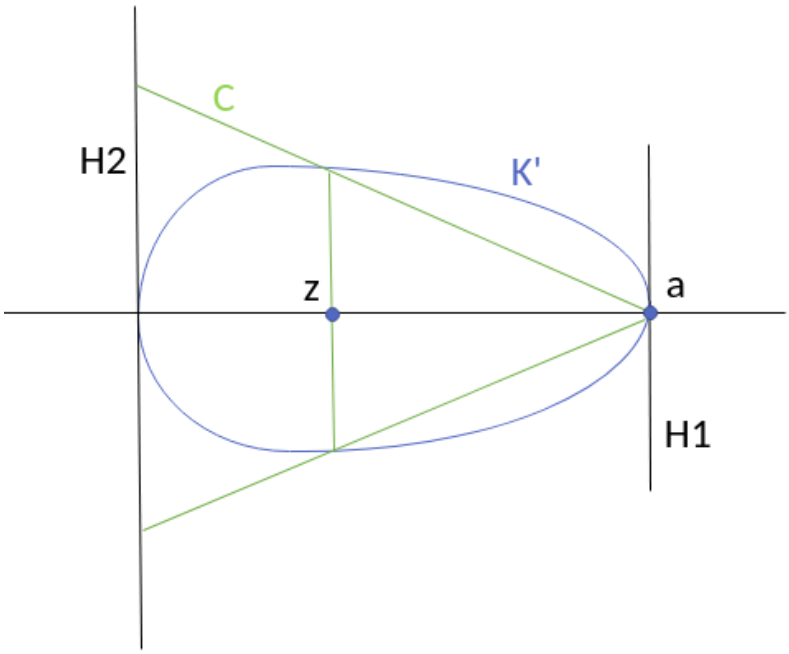
Lemma 3.2.5. *Given body K with centroid at the origin, for any unit vector θ,*

$$\frac{h_K(\theta)}{h_K(-\theta)} \leq n$$

Proof. Again, for simplicity, assume that θ is the positive x_1 direction. Consider replacing each slice $K(t)$ with an $(n-1)$-dimensional ball of equal volume to create a body K'. Let $r(t)$ be the radius of this ball. The volume of $K(t)$ is $f(n-1) \cdot r(t)^{n-1}$. By 3.2.2, $r(c) \geq \lambda r(x) + (1-\lambda) r(y)$. Hence, r is concave.

Let H_1, H_2 be the two supporting hyperplanes of K (and thus of K' as well). Let z be the centroid of K and a the point on H_1 along θ. Consider the cone C' with apex a and base $K(z)$. Extend this to cone C which has base on H_2 as shown below.

Figure 3.2: Cone C and rounded body K' from lemma 2.5



For C, we have precisely $\frac{h_K(\theta)}{h_K(-\theta)} = n$. So, we want to show that z is to the right of c, the centroid of the cone C. This would imply the inequality. Suppose instead that c is to the right of z. Define the following

$$H_z^+ = K' \cap \{x \in \mathbb{R}^n : x_1 \geq z_1\}$$
$$H_z^- = K' \cap \{x \in \mathbb{R}^n : x_1 \leq z_1\}$$
$$H_c^+ = C \ \cap \{x \in \mathbb{R}^n : x_1 \geq c_1\}$$
$$H_c^- = C \ \cap \{x \in \mathbb{R}^n : x_1 \leq c_1\}$$

Since c is the centroid of C, the moments of H_c^+ and H_c^- are equal about the hyperplane through c. Call this value M. By convexity, $H_z^+ \supset H_c^+$, so the momentum, M^+, of H_z^+ about the hyperplane through z is larger than M. Similarly, since the radius function is concave, $H_z^- \subset H_c^-$, hence the momentum, M^-, of H_c^- about the hyperplane through z is less than M. This leads to a contradiction since M^- should equal M^+ as z is the centroid of K'.

□

The proof of 3.2.4 is very similar, so we omit it here. Both of these results are tight, which can be seen for a cone.

Isotropy

Definition 3.2.6 (Isotropic Position)**.** *A set S is in isotropic position if both $\mathbb{E}_S[x] = 0$ and $\mathbb{E}_S[x^T x] = I$.*

Moreover, we for any convex set S, there exists an affine transformation which brings S into isotropic position. We may assume that $\mathbb{E}_S[x] = 0$ since a simple translation can bring the centroid to the origin. Let C be the covariance matrix $\mathbb{E}_S[xx^T]$. Since C is positive definite, there exists a matrix A with $A^2 = C$. So, the body $AS = \{A^{-1}x : x \in S\}$ is isotropic: letting $y = A^{-1}x$, $\mathbb{E}[yy^T] = A^{-1}\mathbb{E}[xx^T]A^{-1} = I$.

Theorem 3.2.7. *For a body $K \subset \mathbb{R}^n$ in isotropic position,*

$$\sqrt{\frac{n+2}{n}}B^n \subseteq K \subseteq \sqrt{n(n+2)}B^n$$

Essentially, a ball of unit radius is contained in K, and a ball of radius $O(n)$ contains K. Isotropy is a useful property that we'll make use of throughout. Often, if we are interested in some affine-invariant property of a body, we can restrict our attention to isotropic bodies. For example, the centroid of a body is affine-invariant. That is, if f is some affine transformation, f maps the centroid of K to the centroid of $f(K)$.

3.2.4 Bounding Iterations

Note that since the center of gravity is invariant under affine transformations, so is the algorithm described. We now focus on bounding m for a body in isotropic position, which will directly imply the general case.

Suppose K contains a cube of side length r, and is contained in a cube of side length R. Recall, the algorithm terminates when the width along each standard direction, e_j is at most r. Thus, for the support $[a, b]$ along e_j we have $a, b \geq r/n$. The final volume of K_m will be $(r/n)^n$. The initial volume is at most R^n. Moreover, by 3.2.4

$$\frac{\text{vol}(K_{i+1})}{\text{vol}(K_i)} \leq 1 - \frac{1}{e}$$

That is, the volume at each step is guaranteed to decrease by a constant fraction. Therefore, $m = O(n \log nR/r)$. For an isotropic body, this comes out to $O(n \log n)$ phases.

However, computing the centroid exactly is #P-hard [4]. On the other hand, the average of $O(\log^2 m)$ sample points provides a good approximation of the centroid if K is a polytope, and $O(n)$ for arbitrary convex bodies [5]. By "good" approximation, we mean that in expectation, any halfspace containing their average will also contain a constant fraction of the volume.

It remains to find these uniform samples. Existing algorithms do this via Markov Chains methods. We now give a brief overview of the main ideas.

3.3 Sampling

A geometric random walk is a sequence of points $x_0, x_1, \ldots$ in $\mathbb{R}^n$ for which x_{i+1} is chosen from a neighborhood of x_i according to some distribution depending only on x_i. In particular, we aim for random walks within K whose stationary distribution is the uniform distribution over K. Additionally, we wish to bound the rate of convergence. Here, we have a brief discussion of random walks, as well as isoperimetry, a key tool used in their analysis.

3.3.1 Examples

- Grid Walk

  The first major breakthrough in a polynomial-time algorithm for approximating volume came from Dyer and Frieze [6] in which they first introduced the ball walk. In it, we define a grid of width δ, and begin at a gridpoint x_0 within the body. With probability $\frac{1}{2}$ we stay in place, and with probability $\frac{1}{2}$ choose point y uniformly at random from the $2n$ neighbors of x_i. If $y \in K$, set $x_{i+1} = y$, otherwise, $x_{i+1} = x_i$.

- Ball Walk

  This is another lazy walk, as with probability $\frac{1}{2}$ we stay in place. Here, we pick a uniform random point y from the ball of radius δ centered at the current point x. If $y \in K$, go to y. Otherwise, stay at x. Some care needs to be taken here. If x is near a corner, it possible that an exponentially small portion of the ball centered at x will lie within K. This would imply that to take a single step, we may require an exponential number of steps.

Intuitively, it seems like this may not be an issue since during the algorithm the probability of reaching such a corner is equally small. However, x_0 itself may be one of these pathological points. One possible solution is to begin from a *warm start*. That is, we have an initial point drawn from a distribution already close to the uniform distribution (closeness here is measured by total variation distance).

We now discuss some main techniques on bounding the rate of convergence of random walks to their stationary distribution. First, we introduce the notion of conductance.

Definition 3.3.1. *The* conductance *of a Markov chain is*

$$\Phi(A) = \min_{0<Q(A)\leq 1/2} \frac{\int_A P_u(K \setminus A) dQ(u)}{Q(A)}$$

where $P_u(X)$ is the probability of taking a step from point u to the set X.

Suppose the conductance of a Markov chain is small, then there must be some subset from which it is hard to escape from. This in turn would require more iterations to converge to the limit distribution. We make this intuition more precise:

Theorem 3.3.2. *We define the "distance" between two distributions P, Q as their total variation distance $\|P - Q\|_{TV}$. Let Q_i be the distribution after i iterations, and Q the limit distribution. Then,*

$$\|Q - Q_i\| \leq \|Q - Q_0\| \cdot \left(1 - \frac{\Phi^2}{2}\right)^i$$

So, it remains to bound the conductance. This notion is closely related to isoperimetry, a purely geometric property of the body.

3.3.2 Isoperimetry

The classical isoperimetric problem is to find a set with minimal surface area for a given volume. The solution to this problem, the sphere, has been known since the ancient Greeks. Here we pose a similar question regarding the boundaries of a surface dividing K into two parts: what is the largest ψ such that

$$\mathrm{vol}_{n-1}(\partial S) \geq \psi \min\{\mathrm{vol}(S), \mathrm{vol}(K \setminus S)\}$$

for any $S \subset K$.

Let's take a closer look at the conductance of the ball walk. To have any chance of taking a step from S to $K \setminus S$, we must initially be near the boundary between the two sets. If the volume of points near the boundary is large then it seems likely that the conductance is high. Indeed,

Theorem 3.3.3 (Isoperimetry [7]). *The ball walk mixes from a warm start in*

$$O^*\left(\frac{n^2}{\psi^2}\right)$$

steps.

This is a fairly broad overview of the main ideas. For a complete survey on geometric random walks, we defer to [8].

3.4 Dynamical Systems

We now change track and discuss dynamical systems which will lead us into dynamical billiards. The next sections present some main ideas that hint toward a different approach to finding samples.

Definition 3.4.1. *A* topological dynamical system *consists of a metric space X and a continuous map $T : X \to X$. T is called* reversible *if it is a homeomorphism. We call X the* phase space *of T.*

Our primary focus will be in describing the trajectory of points under repeated application of the map T. Here, we present the main notions that we will be using in later sections.

Definition 3.4.2. *Given a dynamical system $T : X \to X$, the* orbit *of a point $x \in X$ is the set $\mathcal{O}_x = \{T^k(x) : x \in \mathbb{Z}\}$*

Definition 3.4.3. *A point $x \in X$ is a* fixed *point of T if $T(x) = x$ and therefore, $\mathcal{O}_x = \{x\}$. Similarly, a point $x \in X$ is a* periodic *point of T if there exists some $k \geq 1$ such that $T^k(x) = x$. The minimal such k is the the* period *of x. Moreover, x is a fixed point of T^k.*

Definition 3.4.4. *We say T is a contraction if there exists a constant $0 \leq \tau < 1$ such that*

$$d(T(x), T(y)) \leq \tau d(x, y), \ \forall x, y \in X$$

Theorem 3.4.5 (Banach Fixed-Point Theorem). *Let (X, d) be a complete metric space with contraction mapping $T : X \to X$ with contraction constant τ. Then, T admits a unique fixed point $x^* \in X$. Moreover, for any $x_0 \in X$, the sequence $\{x_n\}$ defined by $x_n = T(x_{n-1})$ converges to x^* at the following rate:*

$$d(x^*, x_n) \leq \frac{\tau^n}{1-\tau} d(x_1, x_0)$$

3.4.1 Hyperbolicity

Hyperbolicity is characterized by a uniform contraction in one direction and a uniform expansion in the other direction. Formally,

Definition 3.4.6. *Let M be a compact C^∞ Riemannian manifold, and let $\phi : M \to M$ be a diffeomorphism. An invariant set $\Lambda \subset X$ is* hyperbolic *if for each $x \in \Lambda$, the tangent space $\mathcal{T}_x M$ splits into a direct sum*

$$\mathcal{T}_x M = E^s(x) \oplus E^u(x)$$

where E^s, E^u are invariant under T in the sense that

$$D\phi(x)E(x) = E(\phi(x)),\ \forall x \in \Lambda$$

for both $E = E^s$ and $E = E^u$. Additionally, there exist constants $C \geq 1$ and $0 < \mu < 1$ so that

$$|D\phi^k(x)v| \leq C\mu^k|v|,\ \ \forall v \in E^s(x), \forall k \geq 0$$
$$|D\phi^{-k}(x)v| \leq C\mu^k|v|,\ \ \forall v \in E^u(x), \forall k \geq 0$$

Hyperbolicity is a particularly useful property of a dynamical system. Many interesting results follow from it, as we see in later sections.

3.5 Dynamical Billiards

A billiard system is defined by a domain $D \subset \mathbb{R}^n$ and the free motion of a point particle inside D with mirror-like reflections off the boundary ∂D. In two dimensions, reflection is most easily described as enforcing that the angle of incidence is equal to the angle of reflection. In higher dimension this still holds true, however there are many directions which would all have the same angle of reflection. Instead, we relate the direction of movement v' after the collision to the direction v before the collision as

$$v' = v - 2\langle v, n\rangle n$$

where n is the unit normal vector to the point of collision. The boundary ∂D is described as a finite union of smooth ($C^k, k \geq 3$) $(n-1)$-dimensional surfaces, $\partial D = \Gamma_1 \cup \cdots \cup \Gamma_r$. For example, if D is a polytope, then Γ_i are its facets. Note that reflection is not properly defined at the "edges" of D, $\partial\Gamma_1 \cup \cdots \cup \partial\Gamma_r$. We distinguish between three types of boundaries.

Definition 3.5.1. Γ_i *is said to be* dispersing *if it is convex inward,* focusing *if it is concave inward, and* flat *otherwise.*

We now present billiards in the setting of dynamical systems. The state of a particle can be described by its position $q \in D$ along with the direction it is moving in, $v \in S^{n-1}$.

Definition 3.5.2 (Phase Space)**.** *The phase space of the system is*

$$\Omega = \{(q, v) : q \in D, v \in S^1\} = D \times S^{n-1}$$

Let $\pi_q(x) = q$ *be the projection of* $x = (q, v) \in \Omega$ *onto* D *and* $\pi_v(x) = v$ *the projection onto* S^{n-1}*.*

We define a *billiard flow* on the phase space as a family of maps $\{\Phi^t\}$. Φ^t advances a point in Ω forward in time by t units. This family forms a group as $\Phi^{s+t} = \Phi^s \circ \Phi^t$ and $\Phi^{-s} \circ \Phi^s = \Phi^0 = \mathrm{Id}$.

Definition 3.5.3. *The* orbit*,* $\mathcal{O}_x$ *of a point* $x \in \Omega$ *is the set* $\{\Phi^t x : t \in \mathbb{R}\}$*. The* billiard trajectory *of* x *is the projection of its orbit,* $\{\pi_q(\Phi^t x)\}$*, onto the table.*

A natural discretization of billiards is to directly map boundary points to boundary points since the trajectory between collisions is a straight line. Define a hypersurface $M \subset \Omega$ which consists of the post-collision vectors after reflection with the boundary: $M = \{(q, v) \in \Omega : q \in \partial Q, \langle v, n\rangle \geq 0\}$. Finally we define a return map $T : M \to M$ which describes successive collisions of the orbit of x. Let $Tx = \Phi^{\tau(x)}x$ where $\tau(x) = \min\{t > 0 : \Phi^t x \in M\}$.

Let's go back to the original volume problem. Suppose that some billiard trajectory will spend time in a subset $A \subset K$ proportional to $\mathrm{vol}(A)/\mathrm{vol}(K)$ for any A. That is, the orbit uniformly covers the polytope. Clearly, we cannot run this process indefinitely, however, perhaps it converges to a near uniform distribution in some polynomial number of reflections which we can use in place of random samples. Billiard tables that have all Γ_i dispersing have some of the strongest properties. Indeed, in two dimensions, if the boundary is additionally smooth, the billiards are hyperbolic, ergodic, mixing [9], and have exponential decay of correlations [10]. We define these in a later section, but roughly they imply a fast rate of convergence to their limit distribution.

3.5.1 Unfolding

For simplicity, we express the following ideas mainly in 2 dimensions, although they easily generalize to higher dimension. Here we present a useful approach to analyzing billiards in polygonal tables. Instead of reflecting the orbit off the edge of the polygon it intersects, we reflect the polygon itself over this edge (see figure 3.3).

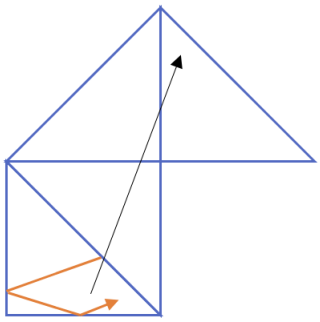


Figure 3.3: Unfolding

If we folded the triangles back, the black line would directly overlap with the true trajectory, the orange line. This allows us to inspect a single straight line. The cleanest unfoldings are for polygons which can edge-tessellate the plane such as the square and the triangle shown above.

3.5.2 The Square

Before taking a look at the general case of billiards, we present some key results for the square as it is simpler to work with. We will look at orbits moving along rational directions and show that these are always periodic. Moreover, orbits of irrational directions are not periodic and uniformly cover the square.

Uniformity

Here, we will assume that v is irrational. We first show that we can reduce this to a one-dimensional problem.

Without loss of generality, assume that the square has side length 1, $v_y = 1$, $v_x \in \mathbb{R} \setminus \mathbb{Q}$, and the initial point q_1 lies on the bottom edge of the square. Let's take a closer look at the tiling created by unfolding the square. Each square can be identified by a row and column as shown below, where $(0, 0)$ corresponds to the initial square.

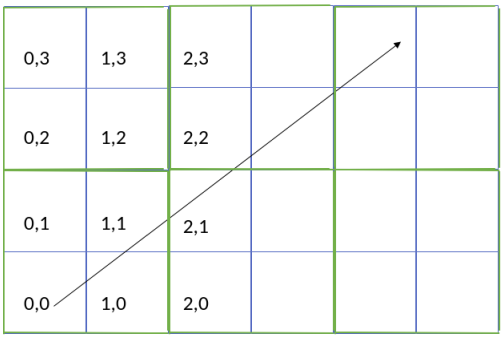


Figure 3.4: Tiling

Squares at coordinates (a_1, b_1), (a_2, b_2) are translations of each other if $a_1 \equiv a_2 \pmod 2$ and $b_1 \equiv b_2 \pmod 2$, since an even number of reflections along the same direction has no change on symmetry. So, the green squares above are translations of each other.

Let q_i denote the i^{th} intersection of the line with the horizontal edges of the green squares. These points map back to points on the bottom edge of our initial green square at $(0, 0)$. Indeed, these will be $((q_i)_x \bmod 2, 0)$. Later, we prove that these points are equidistrbituted along that edge. We first use this to prove the uniformity of the orbit.

For the moment, let's only consider portions of the orbit which travel from the bottom edge to the top edge of the square. These are described by the black arrows in the image below. Since the bottom endpoints are equidistributed, the orbits themselves uniformly cover the space shown.

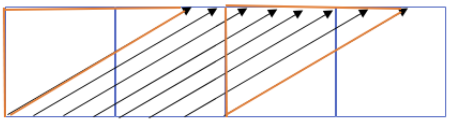


Figure 3.5: Partial Unfolding

The two orange triangles are translations of each other, so indeed the partial orbit we're considering uniformly covers the square. A similar argument holds for the portions of the orbit which begin at the top edge and travel to the bottom edge, so the entire orbit uniformly covers the square.

Lemma 3.5.4. *The sequence* $a_i \pmod 2$ *is equidistributed in the interval* $[0, 2]$.

Proof. The line will intersect the vertical edge after moving 2 units upwards. Therefore, it moves $2v_x$ units to the right before hitting the bottom. So, we are interested in showing that the sequence $a_0 + 2kv_x \pmod 2$, where v_x is irrational, is equidistributed in $[0, 2]$. This is a well known result due to Weyl [11].

□

This proof of uniformity is fairly specific to the square, and is not generalizable to other shapes. In the next section, we explore a more useful criterion for determining uniformity.

3.6 Ergodicity and Mixing

Definition 3.6.1. *Let $T : X \to X$. A function f is T-invariant if $f \circ T = f$. Similarly, a set $A \subset X$ is T-invariant if $T^{-1}(A) = A$.*

Definition 3.6.2. *Let T be a measure-preserving map on a probability space $(X, \mathcal{A}, \mu)$. T is ergodic if every T-invariant set has measure 0 or 1.*

Proposition 3.6.3. *If a measurable function $f : X \to \mathbb{R}$ is invariant under an ergodic map T, then f is constant almost everywhere.*

Proof. Define the level sets $A_c = \{x \in X : f(x) \leq c\}$. We first show that A_c is T-invariant. Suppose $x \in A_c$. Then $f(x) \leq c$, and by invariance, $f(T(x)) \leq c$. Finally $T(x) \in A_c$ and so $A_c \subset T^{-1}(A_c)$. We can similarly show $T^{-1}(A_c) \subset A_c$ and hence $T^{-1}(A_c) = A_c$. By the ergodicity of T, $\mu(A_c) = 0$ or $\mu(A_c) = 1$. Let $p = \inf\{c : \mu(A_c) = 1\}$. Then, since $\mu(A_{p-1/n}) = 0$, $f(x) \geq p$ a.e. and since $\mu(A_p) = 1$, $f(x) \leq p$ a.e. The claim follows.

□

Theorem 3.6.4 (Birkhoff-Khinchin)**.** *Let $(X, \mathcal{A}, \mu)$ be a probability space, and $T : X \to X$ a measure-preserving map. If $f : X \to \mathbb{R}$ is an integrable function, the limit*

$$\tilde{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

exists for almost everywhere $x \in X$, and the function $\tilde{f}$ is T-invariant, integrable, and

$$\int_X \tilde{f} d\mu = \int_X f d\mu$$

The function $\tilde{f}$ is called the time average *of f.*

With this, we investigate how the orbit of a point x under an ergodic map T covers X.

Definition 3.6.5 (Frequency of Return)**.** *Let $A \subset X$, $x \in X$. The limit*

$$\tau(x, A) = \lim_{n\to\infty} \frac{1}{n} \text{card}\{0 \leq m < n : T^m(x) \in A\}$$

is defined as the frequency of returns of the point x to the set A.

Let χ_A denote the indicator function of A. Then,

$$\tau(x, A) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_A(T^j(x))$$

As defined in theorem 3.6.4, with χ_A taking the place of f, $\tau(x, A) = \tilde{\chi}_A(x)$. Therefore,

$$\begin{aligned}\int_X \tau(x, A)d\mu(x) &= \int_X \tilde{\chi}_A(x)d\mu(x) \\ &= \int_X \chi_A(x)d\mu(x) \\ &= \mu(A)\end{aligned}$$

with the second equality following from 3.6.4. Note that we have not yet used the fact that T is ergodic. We can easily see that τ is T-invariant as a function of x since $\tau(x, A) = \tau(T(x), A)$. By 3.6.3, it follows that $\tau(x, A)$ is a constant and so

$$\begin{aligned}\mu(A) &= \int_X \tau(x, A)d\mu(x) \\ &= \tau(x, A) \int_X d\mu(x) \\ &= \tau(x, A)\mu(X) \\ &= \tau(x, A)\end{aligned}$$

for almost everywhere $x \in X$. Hence, the orbit of $x \in X$ almost surely spends time in the set A proportional to its measure, $\mu(A)$. We now present a stronger property than ergodicty: mixing.

Definition 3.6.6. *A measure-preserving endomorphism T on a probability space $(X, \mathcal{A}, \mu)$ is said to be* mixing *if for any $A, B \in \mathcal{A}$,*

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B)$$

Intuitively, one can see this as describing the mixing of two different liquids. Suppose we're pouring red-colored water into a glass of water. Let A be the region originally occupied by the red water and B is any subset of the glass. We expect that eventually the proportion of red water in B will be the same as in the entire glass.

Proposition 3.6.7. *Any mixing map is also ergodic.*

Proof. Suppose $A \subset X$ is a T-invariant set. Since $T^{-n}(A) = A$, $\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A \cap B)$. Since T is mixing, $\mu(A \cap B) = \mu(A)\mu(B)$. Take $A = B$. Then, $\mu(A) = \mu(A)^2$ and hence $\mu(A) = 0$ or $\mu(A) = 1$. □

3.7 Decay of Correlations

Given that a system is ergodic, we would additionally like to describe the rate at which a trajectory approaches its final distribution. Consider the sequence

$$S_n = \sum_{j=0}^{n-1} f(T^j(x))$$

By the Birkhoff Ergodic Theorem 3.6.4, S_n/n converges to $\int_X f d\mu$ almost everywhere. Intuitively, this can be thought of as the law of large numbers in a dynamical systems setting: time average converges to space average. Are there any other relations to probability theory? For instance, does $f(T^n(x))$ effectively become independent of $f(x)$? We make this precise via correlation functions:

Definition 3.7.1. *We denote by C_f the time correlation function of dynamical system T where*

$$C_f(n) = \int_X f(T^n(x))f(x)d\mu - \left(\int_X f(x)d\mu\right)^2$$

Proposition 3.7.2. *Let $T : X \to X$ be a mixing map preserving measure μ. For any $f, g \in \mathcal{L}^2(X)$,*

$$\lim_{n\to\infty} \int_X f(T^n(x))g(x)d\mu = \int_X f(x)d\mu \int_X g(x)d\mu$$

Specifically, if the map T is mixing, C_f converges to 0, and we say that the correlations *decay*. Just as we had an analog of the law of large numbers, we also have one for the central limit theorem.

Definition 3.7.3. *We say that f satisfies the central limit theorem if*

$$\lim_{n\to\infty} \mu\left\{x : \frac{S_n(x) - n\int_X f d\mu}{\sqrt{n}} < z\right\} = \frac{1}{\sqrt{2\pi\sigma_f}} \int_{-\infty}^{z} \exp\frac{-s^2}{2\sigma_f^2} ds$$

for all $z \in \mathbb{R}$, where $\sigma_f \geq 0$ is a constant related to the correlation function as

$$\sigma_f^2 = C_f(0) + 2\sum_{n=1}^{\infty} C_f(n)$$

While the Birkhoff Ergodic theorem assured the existence of the limit, the above provides detail into the rate of convergence. The deviation of S_n/n from $\int_X f d\mu$ (scaled by $1/\sqrt{n}$) is asymptotically Gaussian. However, this only holds if $\sum_{n=1}^{\infty} C_f(n)$ converges. We say this converges *exponentially* if $|C_f(n)| \leq c \cdot e^{-an}$ for some constants $a, c > 0$ and *polynomially* if $|C_f(n)| \leq c \cdot n^{-a}$.

3.8 Shadowing

For the moment, let (X, d) be a metric space, and $T : X \to X$ a dynamical system.

Definition 3.8.1. *A sequence $\xi = \{x_k \in X : k \in \mathbb{N}\}$ is a δ-pseudo-orbit of T if*

$$d(T(x_k), x_{k+1}) \leq \delta, \ \forall k \in \mathbb{N}$$

Definition 3.8.2. *We say that a point $x \in X$ ϵ-shadows a δ-pseudo-orbit $\xi = \{x_k\}$ if*

$$d(T^k(x), x_k) \leq \epsilon, \ \forall k \in \mathbb{N}$$

Additionally, T has the shadowing property *on $Y \subset X$ if given $\epsilon > 0$, there exists $\delta > 0$ such that for any δ-pseudo-orbit $\xi \subset Y$, there is a point x that ϵ-shadows ξ.*

Every pseudo-orbit, which can be thought of as a numerically-computed orbit with rounding errors introduced at each step, stays uniformly close to some true orbit (note this may have a slightly perturbed initial position). That is, the pseudo-trajectory we generate is *shadowed* by a true one.

We are interested in determining which classes of dynamical systems T have the shadowing property, as well as for which sets $Y \subset X$.

Lemma 3.8.3. *Suppose a dynamical system $T : X \to X$ is a contraction with constant τ. Then T has the shadowing property.*

Proof. Let $\mathbf{x} = \{x_n\}_{n\in\mathbb{Z}} \subset X$ be a δ-pseudo-orbit. Define the metric space $(\mathbf{E}, D)$ by

$$\mathbf{E} = \{\mathbf{y} : \mathbf{y} = \{y_n\}_{n\in\mathbb{Z}}, d(x_n, y_n) \leq \epsilon\}$$

with metric

$$D(\mathbf{x}, \mathbf{y}) = \sup\{d(x_n, y_n) : n \in \mathbb{Z}\}$$

That is, $\mathbf{E}$ contains all sequences which are pointwise within ϵ of our δ-pseudo-orbit, $\mathbf{x}$. Define a sequence $\mathbf{T}(\mathbf{y})$ for sequences $\mathbf{y}$ as

$$\mathbf{T}(\mathbf{y})_n = T(y_{n-1})$$

If $\mathbf{E}$ contains an orbit of T, the claim follows. We aim to show the existence of a fixed point of $\mathbf{T}$, since such a point would be an orbit of T. We first prove $\mathbf{E}$ is $\mathbf{T}$-invariant. By the triangle inequality,

$$d(T(y_{n-1}), x_n) \leq d(T(y_{n-1}), T(x_{n-1})) + d(T(x_{n-1}), x_n)$$

Since T is a contraction, and $\{x_n\}$ is a δ-pseudo-orbit,

$$d(T(y_{n-1}), x_n) \leq \tau d(y_{n-1}, x_{n-1}) + \delta$$

We can choose our value of δ for any given ϵ. Note that $d(y_{n-1}, x_{n-1}) \leq \epsilon$ since $\mathbf{y} \in \mathbf{E}$. Letting $\delta = (1-\tau)\epsilon$,

$$\begin{aligned} d(T(y_{n-1}), x_n) &\leq \tau\epsilon + (1-\tau)\epsilon \\ &\leq \epsilon \end{aligned}$$

Hence, for any $\mathbf{y} \in \mathbf{E}$, $\mathbf{T}(\mathbf{y}) \in \mathbf{E}$ and so $\mathbf{E}$ is indeed $\mathbf{T}$-invariant. Finally, we show that $\mathbf{T}$ is a contraction with contraction constant τ. Take $\mathbf{y}, \mathbf{z} \in \mathbf{E}$. Then,

$$\begin{aligned} D(\mathbf{T}(\mathbf{y}), \mathbf{T}(\mathbf{z})) &= \sup\{d(T(y_{n-1}), T(z_{n-1})) : n \in \mathbb{Z}\} \\ &\leq \tau \cdot \sup\{d(y_{n-1}, z_{n-1}) : n \in \mathbb{Z}\} \\ &\leq \tau D(\mathbf{y}, \mathbf{z}) \end{aligned}$$

By the Banach Fixed-Point theorem 3.4.5, $\mathbf{T}$ has a fixed point, hence proving the claim. □

Shadowing holds with greater generality. As hyperbolicity was a key property for many of the previous properties, it is for shadowing as well. Note that in the previous lemma, T was indeed hyperbolic as $E^u = 0$.

Lemma 3.8.4. *If Λ is a hyperbolic set for a diffeomorphism ϕ, then there exists a neighborhood W of Λ such that ϕ has the shadowing property on W.*

Bibliography

[1] M. Dyer and A. Frieze. A random polynomial time algorithm for approximating the volume of convex bodies. In *Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing*, STOC '89, pages 375–381, New York, NY, USA, 1989. ACM.

[2] László Lovász and Santosh Vempala. Simulated annealing in convex bodies and an o*(n4) volume algorithm. *J. Comput. Syst. Sci.*, 72(2):392–417, March 2006.

[3] Imre Bárány and Zoltán Füredi. Computing the volume is difficult. *Discrete & Computational Geometry*, 2(4):319–326, 1987.

[4] Luis A. Rademacher. Approximating the centroid is hard. In *Proceedings of the Twenty-third Annual Symposium on Computational Geometry*, SCG '07, pages 302–305, New York, NY, USA, 2007. ACM.

[5] Dimitris Bertsimas and Santosh Vempala. Solving convex programs by random walks. *J. ACM*, 51(4):540–556, 2004.

[6] Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, January 1991.

[7] Ravindran Kannan, László Lovász, and Miklós Simonovits. Random walks and an $o^*(n^5)$ volume algorithm for convex bodies. *Random Structures and Algorithms - RSA*, 11:1–50, 08 1997.

[8] Santosh Vempala. Geometric random walks: a survey. *Combinatorial and computational geometry*, 52(573-612):2, 2005.

[9] N. Chernov, R. Markarian, and American Mathematical Society. *Chaotic Billiards*. Mathematical surveys and monographs. American Mathematical Society, 2006.

[10] N. Chernov. Decay of correlations and dispersing billiards. *Journal of Statistical Physics*, 94(3):513–556, Feb 1999.

[11] Hermann Weyl. Über die gleichverteilung von zahlen mod. eins. *Mathematische Annalen*, 77(3):313–352, Sep 1916.