Random Billiards
Nicole Andrews
Spring Semester of 2019

1 Introduction

A billiard dynamical system is a mathematical model in which a point particle, perhaps modeling an infinitesimally small gas molecule, is constrained to move within an environment, called the billiard table, interacting with the boundary of the environment through collisions.

In particular, we study random billiards. Random billiards are defined as billiard models whose collision law is specified by assigning the post-collision velocity vector to be random, distributed according to a certain probability distribution.

In this paper, we first consider a scenario in which a particle is constrained to a table that is tubular in shape. We simulate the model and plot the trajectory and we also examine restricting the length of the tube and examine the positions of the particle and their respective frequencies. Next, we consider a similar scenario, but for a particle that has some internal state of either hot or cold which is governed by a Hidden Markov Model. In part 3 of this paper, we consider a particle that is constrained inside of a unit disk table. Finally, there is a mathematical appendix at the end of this paper which defines important mathematical terminology while also attempting to explain them in the context of this project.

2 Part 1

The initial model we consider is that of a basic tubular random billiard table. Consider a table $Q$ that is in the shape of a tube. Let $Q$ be that tubular table. We assume that $Q$ consists of a boundary defined by the horizontal lines $y = 0$ and $y = 1$ and the region in between those two horizontal lines. Assume there is a particle contained in this table. The particle moves with constant, unit speed, velocity in between collisions. Let $\Theta \sim \text{Unif}(\Theta_1, \Theta_2, \Theta_3,...)$ be a random variable taking on four independent and identically distributed angles from the set $(0, \pi)$. These angles represent the angle between the trajectory’s velocity vector and the boundary curve after each collision. Whenever the particle collides with the billiard table, we sample uniformly at random from $\Theta$. If the collision was with the bottom boundary component of the table $y = 0$, we
set the post-collision velocity vector to be $V = (\cos \theta, \sin \theta)$. If the collision was with the top boundary component of the table $y = 1$, we also sample uniformly at random from $\Theta$, however this time we set the post-collision velocity vector to be $V = (\cos \theta, -\sin \theta)$. In either case, after the new velocity vector is computed, the particle will then move with the speed and direction indicated by the new velocity until it again collides with the boundary of the table and the process described above is repeated again.

We simulated the above model and ran our simulation for 100 collisions, letting the initial position be at the origin $(0, 0)$. A plot of the trajectory of the particle can be seen below in Figure 1.

![Figure 1: A particle is contained within the tube defined by the horizontal lines $y = 0$ and $y = 1$. The particle starts at the origin $(0, 0)$ and continues with unit speed velocity until it collides with either horizontal line in which case an angle is chosen uniformly at random and a new post-collision velocity vector is found based on the angle that was chosen. For this specific simulation, we let $\Theta \sim \text{Unif}\{\frac{\pi}{5}, \frac{2\pi}{5}, \frac{3\pi}{5}, \frac{4\pi}{5}\}$.

Recall that each angle is independent from all of the other angles and identically distributed to all of the other angles. Let $\Theta_k$ be the $k$th angle chosen uniformly at random. Then it clearly follows that $\Theta_k$ is i.i.d. Let $Z_k$ be the horizontal distance between the particle’s position when it collided with the boundary and the position of the particle upon its next collision. Using ele-
mentary trigonometry, we compute that $Z_k = \cot \Theta_k$. Since the cot of each angle represents a horizontal distance the particle moved, if we want to know the particle’s horizontal position in the tube after some number of collisions, we can just compute the sum of all of the horizontal distances moved. In other words, if we let $S_n$ denote the horizontal position of the particle, then we know that $S_n = \sum_{k=1}^{n} Z_k$, assuming $S_0 = 0$, which was one of our initial assumptions.

Let the expected horizontal position after the $k$th collision be denoted $\mu$. Let the variance of the horizontal position after the $k$th collision be denoted $\sigma^2$.

So, $\mu = E[Z_k]$ and $\sigma^2 = V(Z_k)$. By the Central Limit Theorem, $\frac{S_n - n\mu}{\sigma \sqrt{n}}$ approaches $N(0, 1)$ as $n$ approaches $\infty$. In other words, $\frac{S_n - n\mu}{\sigma \sqrt{n}} \approx N(0, 1)$ if $n$ is large enough.

So, it follows that $S_n - n\mu \approx \sigma \sqrt{n} N(0, 1)$. It further follows that $S_n \approx n\mu + \sigma \sqrt{n} N(0, 1)$. Since we know that $V(aX) = a^2 V(X)$, we have that $S_n \approx n\mu + N(0, \sigma^2 n)$. So, we have that $S_n \approx N(0, \sigma^2 n)$. So, using the Central Limit Theorem, we have shown that the position of the particle after $n$ collisions with the billiard table has an approximately normal distribution based on the variance of $Z_k$ and the number of collisions. See the appendix for definitions and explanations of expected value, variance, and the Central Limit Theorem.

To confirm our expectations for what would happen based on this mathematical theory, we ran the tubular billiard table simulation for 10,000 trials, such that the particle collided with the boundary of the billiard table first 2,800 times, then 3,500 times, then 4,200 times, and finally 5,600 times. We plotted a histogram of the position for each of the simulations. From the histograms, we observe that our expectations were correct. As the number of collisions $n$ increases, the histogram begins to look more normally distributed.

The histograms and Q-Q plots can be seen below.
The above histograms show the position of the particle after a certain number of collisions (2800, 3500, 4200, and 5600, respectively) for 10,000 trials of each simulation. The histograms progressively become more normally distributed as the number of collisions $$n$$ gets larger, which is what we expect based on the Central Limit Theorem. The QQ-Plot demonstrates the normality of the histogram of the position of the particle after 5600 collisions because of the linearity of the points.
Another scenario we are interested in exploring is the scenario in which the length of the tubular billiard table is restricted. We consider a scenario in which the tube extends 300 units to the right and 300 units to the left, such that the total length of the tube is 600 units. In this scenario, we are interested in knowing the y-coordinate of the position of the particle when it first escapes the tube (either escaping out of the left side of the tube or out of the right side of the tube). We run the simulation 50,000 times and keep track of the Y position of the particle upon its escape from the tube for each run of the simulation. We then plot a histogram of the y-values of escape versus the frequency of observing that y-value. The histogram can be seen below, in Figure 6.

![Histogram of Y Position Upon Escape](image)

Figure 7: This histogram relates the y-coordinate of the particle’s position upon escape from the tube with the frequency of the particle being at that y-coordinate when it escapes the tube based on 50,000 runs of the simulation. This histogram appears to be uniformly distributed, which is what we expect given that the angles that are used to calculate the post-collision velocity vectors are chosen uniformly at random.

Another thing we are interested in is how the mean escape time changes as the length of the tube increases. We consider L to be the length to the right from 0 of the tube. We consider L to be 100, then 300, then 500, then 700, then 900, and finally 1100, and for each of these lengths we run the particle simulation 1000 times and calculate the expected escape time. Then we make a plot of the mean escape time versus the positive length.
3 Part 2

For another model that we study, we consider a random billiard model in which there is a Hidden Markov Model that decides the kinds of random reflections that are possible.

To more fully understand this random billiard model, it is important to first solidify what a Hidden Markov Model is. Hidden Markov Models are Markov models in which the system being modeled is assumed to be a Markov model with hidden states. The probability of observing some behavior, which is visible, is determined by the hidden state that the model happens to be in. Similar to Markov Chains, the probability of being in a certain state depends only on the previous state. Hidden Markov Chains have two different matrices associated with them: a state transition matrix and an emission matrix. The state transition matrix illustrates the probabilities of going from all of the different states to all of the different states and the emission matrix illustrates the probability of some behavior occurring given a certain hidden state.

To better clarify this definition and explanation, we can look to our ran-
dom billiard simulation as an example. For this random billiard simulation, we consider a particle that has two possible internal states. The particle is either in a hot state or it is in a cold state. Initially, the particle has a 50% chance of being in the hot state and a 50% chance of being in the cold state. If the particle is in the hot state, then it stays in the hot state with probability $p$ and transitions to the cold state with probability $1 - p$. If the particle is in the cold state, then it stays in the cold state with probability $q$ and transitions to the hot state with probability $1 - q$. These are the transition probabilities of the Hidden Markov Model underlying this billiard model. However, it is also important to think about what behavior can occur and with what probabilities that behavior occurs at each state. We assume that the particle can exhibit two potential behaviors. It either scatters with a high degree of randomness (in 10 different possible directions) or it scatters with a low degree of randomness (in 2 possible directions). If the particle is in the hot state, it scatters with a high degree of randomness with probability $r_1$ and with a low degree of randomness with probability $d_1$. If the particle is in the cold state, it scatters with a high degree of randomness with probability $r_2$ and with a low degree of randomness with probability $d_2$. Looking at the transition matrix and the emission matrix illustrates these probabilities more clearly.

$$\text{transition matrix} = \begin{pmatrix} \text{hot} & \text{cold} \\ \text{hot} & p & 1 - p \\ \text{cold} & 1 - q & q \end{pmatrix}$$

$$\text{emission matrix} = \begin{pmatrix} \text{high degree} & \text{low degree} \\ \text{hot} & r_1 & d_1 \\ \text{cold} & r_2 & d_2 \end{pmatrix}$$

We define the 10 different possible directions for scattering with a high degree of randomness to be $\{\Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5, \Theta_6, \Theta_7, \Theta_8, \Theta_9, \Theta_{10}\}$. We define the 2 different possible directions for scattering with a low degree of randomness to be $\{\Theta_{11}, \Theta_{12}\}$.

A plot of a simulation consisting of a particle governed by a HMM that was run for 100 collisions is illustrated below in Figure 8.
Figure 9: This figure illustrates the trajectory of a particle in a tubular channel that is governed by a Hidden Markov Model. In this case, the probabilities are $p = 0.6$, $q = 0.4$, $r_1 = 0.8$, $d_1 = 0.2$, $r_2 = 0.2$, and $d_2 = 0.8$. If the particle is in the hot state, an angle is chosen from the list $\{\pi/11, 2\pi/11, 3\pi/11, 4\pi/11, 5\pi/11, 6\pi/11, 7\pi/11, 8\pi/11, 9\pi/11, 10\pi/11\}$ uniformly at random. If the particle is in the cold state, an angle is chosen from the list $\{\pi/3, 2\pi/3\}$ uniformly at random. In either case, the angle is used to compute the post-collision velocity vector (this is similar to the non-HMM case where the particle did not have an internal state).

We are also interested in the particle’s position after many steps for a large number of trials. So similarly as for the case in Part 1, we run 10,000 trials where the particle collides with the boundary 2,800 times, then 3,500 times, then 4,200 times, and finally 5,600 times for each trial, and plot histograms of the particle’s position. The histograms below illustrate this.
The above histograms show the position of the particle after a certain number of collisions (2800, 3500, 4200, and 5600, respectively) for 10,000 trials of each simulation. Again, we observe that the histograms progressively become more normally distributed as the number of collisions $n$ gets larger, which is what we expect based on the Central Limit Theorem.
Similarly to the non-HMM tubular random billiard model, we are again interested in observing the y-coordinates of the particle’s position upon its escape and their respective frequencies when we restrict the length of the table. Again letting \(-L\) be \(-300\) and letting \(L\) be \(300\), such that the total length of the tubular channel is \(600\), we again run the simulation 50,000 times and plot a histogram of the Y position of the particle versus the frequency of the particle being in that Y position upon its escape from the tube (Figure 9).

![Figure 15: Again for the HMM case, we are interested in plotting a histogram that relates the y-coordinate of the particle upon its escape from the tubular table with the frequency of the particle being at that y-coordinate when it escapes. We again observe a fairly uniform distribution in the case where the particle is governed by a Hidden Markov Model.](image)

Recall our two-state transition matrix which is illustrated again below.

\[
\begin{pmatrix}
\text{hot} & \text{cold} \\
\text{hot} & p & 1 - p \\
\text{cold} & 1 - q & q
\end{pmatrix}
\]

We assume that \(Y_n\) is the sequence of internal hot or cold states of the particle. We want to write a general expression for the limiting distribution of \(Y_n\). Let \(\lambda\) be the stationary distribution of \(Y_n\). Since our transition matrix is a two-state Markov Chain, we can find the stationary distribution explicitly. Computing
this yields $\lambda = \left[ \frac{1-q}{2(p+q)}, \frac{1-p}{2(p+q)} \right]$. We know that $\Theta_n$ is the post-collision angle the velocity vector takes on after $n$ collisions. We can see that the possible values that $\Theta_n$ can take on, unconditionally (i.e., without knowing the internal hot or cold state of the particle) are $\Theta_n = \{\Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5, \Theta_6, \Theta_7, \Theta_8, \Theta_9, \Theta_{10}, \Theta_1', \Theta_2'\}$.

Now also suppose that the stationary distribution $\lambda$ of $Y_n$ is the initial distribution of the Markov Chain such that $Y_0 \sim \lambda$. We want to find the distribution $\pi_k$ (i.e. the probability mass function) of $\Theta_n$. Using the law of total probability, conditioning on the internal state of the particle at the $n$th collision, we have

$$\pi_k = P(\Theta_n = k) = P(\Theta_n = k|Y_n = \text{hot})P(Y_n = \text{hot}) + P(\Theta_n = k|Y_n = \text{cold})P(Y_n = \text{cold}).$$

Now first suppose that $k = \Theta_j$ for $j = 1, \ldots, n$. Then we know that $P(\Theta_n = k|Y_n = \text{cold}) = 0$, so it follows that $P(\Theta_n = k|Y_n = \text{cold})P(Y_n = \text{cold}) = 0$. So, it further follows that

$$\pi_k = P(\Theta_n = k) = P(\Theta_n = k|Y_n = \text{hot})P(Y_n = \text{hot}) = (0.1)P(Y_n = \text{hot}).$$

Since $Y_n \sim \lambda$ and $Y_0 \sim \lambda$ by initial assumption, it follows that

$$\pi_k = (0.1)P(Y_0 = \text{hot}) = 0.1\lambda_{\text{hot}}.$$

Similarly, we consider the case where $k = \Theta_1'$ or $k = \Theta_2'$. In this case, $P(\Theta_n = k|Y_n = \text{hot}) = 0$, so it follows that $P(\Theta_n = k|Y_n = \text{hot})P(Y_n = \text{hot}) = 0$. So, it further follows that

$$\pi_k = P(\Theta_n = k) = P(\Theta_n = k|Y_n = \text{cold})P(Y_n = \text{cold}) = (0.5)P(Y_n = \text{cold}) = 0.5\lambda_{\text{cold}}.$$

In summary, we have that

$$\pi_k = P(\Theta_n = k) = \begin{cases} 0.5\lambda_{\text{cold}} & \Theta_n = \Theta_1', \Theta_2' \\ 0.1\lambda_{\text{hot}} & \Theta_n = \Theta_1, \ldots, \Theta_n. \end{cases}$$

From this, it is clear that if we let $Y_0 \sim \alpha$, then $Y_n \sim \alpha P^n$ if $\alpha$ is stationary, by definition of a stationary distribution. Then, we have
\[ Y_n \sim \alpha P^n = (\alpha P)P^{n-1} \]
\[ = \alpha P^{n-1} \]
\[ = \alpha P \star P^{n-2} \]
\[ = \alpha P^{n-2} \ldots \]
\[ = \alpha P \]
\[ = \alpha. \]

If we suppose that \( k = \Theta_j \), where \( j = 1, \ldots, 10 \), and \( Y_0 \sim \alpha \), but \( Y_0 \) is not necessarily stationary initially, then we know that \( P(\Theta_n = k) = 0.1P(Y_n = \text{hot}) \). Taking the limit as \( n \) approaches \( \infty \) of both sides yields that the limiting distribution for \( \Theta_n \) is

\[ \lim_{n \to \infty} P(\Theta_n = \Theta_j) = \lim_{n \to \infty} 0.1P(Y_n = \text{hot}) \]
\[ = 0.1 \lim_{n \to \infty} P(Y_n = \text{hot}) \]
\[ = 0.1\lambda_{\text{hot}} \]

Now instead suppose that \( k = \Theta'_1 \) or \( k = \Theta'_2 \). In this case, we have \( P(\Theta_n = \Theta'_j) = 0.5P(Y_n = \text{cold}) \).

Taking the limit as \( n \) approaches \( \infty \), we see that

\[ \lim_{n \to \infty} P(\Theta_n = \Theta'_j) = 0.5\lambda_{\text{cold}} \]

In both of these cases, we can see that the stationary distribution of the transition matrix of this Hidden Markov Model is the limiting distribution, even when the initial state distribution is \( \text{not} \) the stationary distribution.

### 4 Part 3

Another random billiard model we consider is a random billiard model in which the table \( Q \) consists of the unit disk, centered at the origin. This random model is similar to the non-HMM tubular random billiard model, except for the fact that the billiard table is the unit disk centered at the origin rather than a tube and the velocity vector is computed the same way no matter which part of the boundary of the table the particle collides with.

We let the particle move instead of the unit circle until it has collided with the border of the table 700 times. We plot the trajectory, as seen below in Figure 10.
Figure 16: A simulation is made of a particle inside of a unit circle billiard table. The particle moves inside of the table with constant, unit speed velocity until it collides with the boundary of the table. When it collides with the table, the post-collision velocity vector is computed and then the particle moves again, now with the new speed and direction. This simulation was run for 700 collisions. It is possible that this simulation suffers from round-off error, because we expect that the particle can go back to the point it was at previously from the point where it currently is.

We are interested in the points on the unit circle that are visited, so we make a plot of the angle $\phi$ versus the frequency of that angle after running the simulation for 50,000 collisions.
Consider that the particle starts at some position on the unit circle billiard which makes an angle with the unit circle that we denote $\varphi_0$. A value for $\theta$ is chosen uniformly at random, and the particle travels in the table until it reaches its new position which makes a new angle with the unit circle, denoted $\varphi_1$. We can calculate $\varphi_1$ from the previous angle $\varphi_0$ using the following formula: $\varphi_1 = \varphi_0 + X_1$ where $X_1 = -(\pi + 2\theta_1)$. Similarly, for $\varphi_2$ we have $\varphi_2 = \varphi_1 + X_2$ where $X_2 = -(\pi + 2\theta_2)$. Substituting in for $\varphi_1$ yields $\varphi_2 = \varphi_0 + X_1 + X_2$. Continuing this pattern, we have that $\varphi_3 = \varphi_2 + X_3$ and thus that $\varphi_3 = \varphi_0 + X_1 + X_2 + X_3$.

In general, $\varphi_n = \varphi_0 + \sum_{k=1}^{n} X_k$.

We are interested in understanding the trajectory of the particle as it moves around and collides with the boundary of the unit circle billiard table. A plot of this can be seen below in Figure 13.
Figure 18: A simulation is made of a particle inside of a unit circle billiard table. The particle moves inside of the table with constant, unit speed velocity until it collides with the boundary of the table. When it collides with the table, the post-collision velocity vector is computed and then the particle moves again, now with the new speed and direction. This simulation was run for 10,000 collisions. This plot of the trajectory of the particle is more in line with our expectations for the movement of the particle, because we expect that at any given position, the particle has the ability to move back to the position it was in at its previous step.

We are also interested in understanding what the distribution of $\phi_n$ is as $n$ gets large. In other words, we want to know the value of $\lim_{n \to \infty} \phi_n$. This limit is equivalent to $\sum_{k=1}^{\infty} X_k \pmod{2\pi}$. To calculate this, we generate $X_1, ..., X_n$ 10,000 times and calculate $S_n = \sum_{k=1}^{n} X_k \pmod{2\pi}$.

More specifically, in our simulation, the position of the particle after 10,000 collisions, $S_{10,000} = 5.026548245742958$.

Let $\mu = E[X_i]$ and let $\sigma^2 = V(X_i)$. Then $\frac{S_n - n\mu}{\sigma\sqrt{n}} \approx N(0, 1)$. So, $S_n \approx n\mu + N(0, \sigma^2 n)$, so $S_n \approx N(n\mu, \sigma^2 n)$. However, in this case, we want to consider $S_n \pmod{2\pi}$. In other words, we want to think about what the distribution of a normal random variable, mod $2\pi$ is.

We plot a histogram of the angle made with the unit circle for $n = 10,000$ collisions and for 5000 samples. The histogram is presented below in Figure 19.
5 Mathematical Appendix

5.1 Expected Value

$E[X]$ is a weighted average of $X$. The expected value of $X$ can be summarized with the following formula:

$$E[X] = \sum_{k \in S} kP(X = k)$$

where $S$ is the set of values $X$ can take on. As a way of approximating $E[X]$ we often take an i.i.d sample $X_1, X_2, ..., X_n$ of random variables distributed like $X$. Then, the Law of Large numbers says that

$$\left( \frac{X_1 + ... + X_n}{n} \right) \approx E[X].$$

5.2 Variance

Variance is a way to measure variability of a random variable from the mean. Suppose $X$ is a such a random variable with mean $E[X] = \mu < \infty$.

The variance of $X$ is defined as $V[X] = E[(X - \mu)^2] = \sum_x (x - \mu)^2 P(X = x)$. In this formula, the difference $(x - \mu)$ is a description of the deviation of some
particular outcome from the mean. Like expected value, variance is also a weighted average, however it is a weighted average of the squared deviations from the mean. In the case of this project, when we refer to the variance of $X$, we are referring to the variance of some particular position the particle was in after some number of collisions, meaning we are referring to the weighted average of squared deviations from the mean position (the position we expect the particle to be in).

### 5.3 The Law of Large Numbers

Recall that $Z_1, Z_2, ..., Z_n$ are an i.i.d sequence of random variables, where $Z_k$ represents the position of the particle after the $k$th collision. These random variables have finite expectation $\mu$. Also recall that $S_n = X_1 + ... + X_n$ represents the $x$ position of the particle after $n$ collisions. The law of large numbers says that the average $S_n/n$ converges to $\mu$ as $n$ approaches $\infty$. In other words, this means that as $n$ gets larger, we expect the average position of the particle after $n$ collisions that we compute to be approaching the expected position of the particle after $n$ collisions.

### 5.4 The Central Limit Theorem

Broadly, the Central Limit Theorem tells us that the distribution of a standardized random variable $(S_n/n - \mu)/(\sigma/\sqrt{n})$ approaches a normal distribution as $n$ approaches $\infty$. As previously mentioned, in the case of this project, we expect that as the number of collisions $n$ gets larger, the histograms of the final position of the particle will approach normal distributions.