

WIND-TREE MODEL FOR BILLIARD MOTION FROM A SIGNAL PROCESSING VIEWPOINT

ENRICO AU-YEUNG AND NICK KREISSLER

ABSTRACT. In the Ehrenfest wind tree model, a point particle moves on the plane and collides with randomly placed fixed square obstacles under the usual law of geometric optics. The particle represents the wind and the squares are the trees. We examine the periodic version of the model. Previous authors analyze the dynamical properties of the model using techniques from algebraic topology or ergodic theory. In contrast to these works, we adopt a signal processing viewpoint. We describe the phenomenon of the long-term trajectories by using a 3-state hidden Markov model.

1. INTRODUCTION AND MOTIVATION

In the wind tree model introduced by P. Erhenfest and T. Ehrenfest in 1912, a point particle moves on the plane and collides with randomly placed fixed square obstacles under the usual law of geometric optics. The particle represents the wind and the squares are the trees. We investigate the periodic version of the model, where identical square obstacles are placed periodically in the plane. Our aim is to understand some of its dynamical properties (see Figure 1 for the position of the billiard after 15 collisions with the obstacles.)

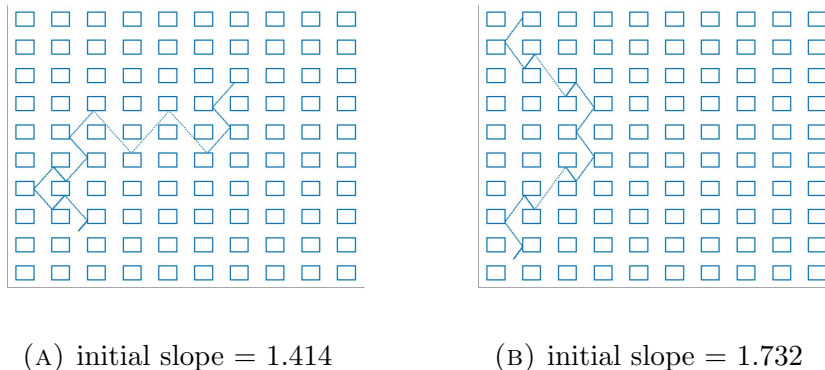


FIGURE 1. Trajectories after 15 collisions

Previous authors analyze the dynamical properties of the model using techniques from algebraic topology (see [1], [2], [3], [4]). In this paper we take a somewhat different point of view than these works. We assume no knowledge of algebraic topology. Instead, we take a signal processing viewpoint and use a hidden Markov model with three hidden states to study the phenomenon we observe.

The billiard trajectory is characterized by a direction at the initial position. This direction is specified by the tangent of the angle with a horizontal line parallel to the plane. The slope m of the initial velocity vector corresponds to the angle θ , i.e. $m = \tan \theta$. Henceforth, we can refer to either the direction θ or the initial slope m .

2. BACKGROUND AND RELATED WORKS

In the periodic version of the Erhenfest wind tree model, due to J. Hardy and J. Weber ([10]), the obstacles are identical rectangular blocks located periodically in the plane, every obstacle centered at each point of \mathbb{Z}^2 . The blocks are rectangles of size $a \times b$, with $0 < a < 1, 0 < b < 1$. We denote by $T(a, b)$ the subset of the plane obtained by removing the obstacles and name its billiard the wind tree model. Let $\phi_t^\theta: T(a, b) \rightarrow T(a, b)$ be the billiard flow: for a point $p \in T(a, b)$, the point $\phi_t^\theta(p)$ is the position of a particle after time t starting from position p in direction θ . Let d be the Euclidean distance in \mathbb{R}^2 . The flow in direction θ is recurrent, if for almost all points x in $T(a, b)$, we have $\liminf_{t \rightarrow \infty} d(x, \phi_t^\theta(x)) = 0$. The flow in direction θ is divergent, if for almost all points x in $T(a, b)$, we have $\liminf_{t \rightarrow \infty} d(x, \phi_t^\theta(x)) = \infty$. Delecroix [3] proves the following result about the set of parameters (a, b, θ) for which the flow in $T(a, b)$ in direction θ is divergent.

Theorem 1. If a and b are either rational or quadratic of the form $1/(1-a) = x+y\sqrt{D}$ and $1/(1-b) = (1-x) + y\sqrt{D}$ there exists a dense set $\Lambda \in [0, 2\pi)$ of Hausdorff dimension not smaller than $1/2$ such that for every $\theta \in \Lambda$ and every point x in $T(a, b)$ with infinite forward orbit $\liminf_{t \rightarrow \infty} d(x, \phi_t^\theta(x)) = \infty$. In particular the flow ϕ_t^θ is divergent.

Delecroix, Hubert, and Lelièvre ([2]) determine the rate of diffusion of the orbits which is valid for almost all direction θ .

Theorem 2. Then for all parameters $(a, b) \in (0, 1)^2$, Lebesgue-almost all θ and every point p in $T(a, b)$ (with an infinite forward orbit)

$$\limsup_{T \rightarrow \infty} \frac{\log d(p, \phi_t^\theta(p))}{\log T} = \frac{2}{3}.$$

Remark: The implication is significant. If we change the height and the width of the obstacle, we can get different billiard trajectories, but this does not change the diffusion rate. To understand the dynamical properties of the billiard motion, it is enough to take all the obstacles to be square blocks.

This theorem is impressive because it is valid for almost all direction θ (in the sense of Lebesgue measure). The set of rational numbers Q is a set of Lebesgue measure 0, therefore, if $\tan \theta \in Q$, we cannot conclude from the theorem whether the particle with direction θ is recurrent or divergent.

We adopt the following setting in the remainder of this manuscript:

- (1) All obstacles are square blocks of side length 1.
- (2) The particle starts at the origin.
- (3) The centers of the blocks are the ordered pairs of all the odd integers.

This setting implies that the first obstacle to the northeast of the origin has its bottom left corner at $(0.5, 0.5)$ and its bottom right corner at $(1.5, 0.5)$ with its center at $(1, 1)$. Two adjacent square blocks are at distance one apart. There is no loss of generality in choosing this specific setting (see the **Remark**).

Sections 3 and 4 contain illustrations for the position of the particle after N collisions with the square blocks. Section 5 shows two experiments for billiard trajectories after 500 collisions, from which we can identify some patterns. These simulations suggest that there are three types of motions. Section 6 reviews some important definitions and notations in hidden Markov processes, also known as hidden Markov models. Section 7 describes how a 3-state hidden Markov process is adequate to model a dynamical property of the periodic wind-tree model.

3. ILLUSTRATIONS: PART 1

Initial slope = 1.414

These figures show the position of the particle after N collisions with square blocks.

Motion of billiard after 50 collisions, after 100, 150, and 300 collisions.

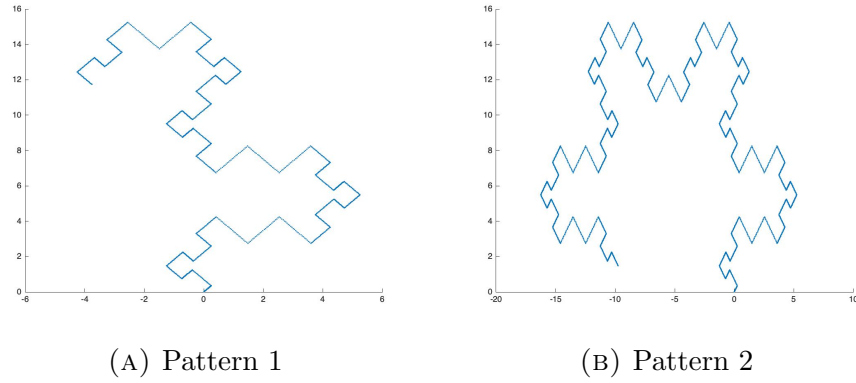


FIGURE 2. Patterns after 50 and 100 collisions

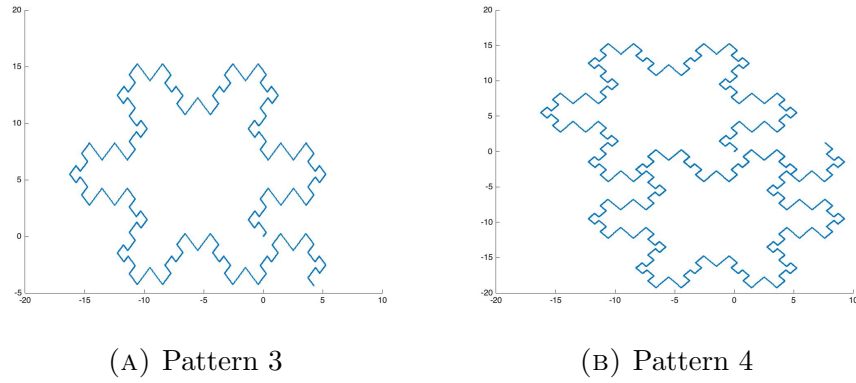


FIGURE 3. Patterns after 150 and 300 collisions

4. ILLUSTRATIONS: PART 2

Initial slope = 1.732

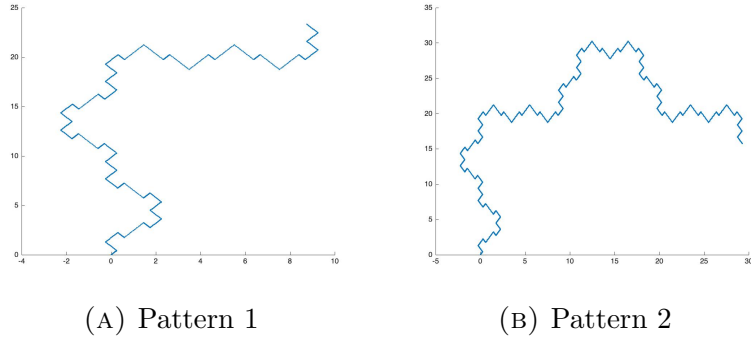


FIGURE 4. Patterns after 50 and 100 collisions

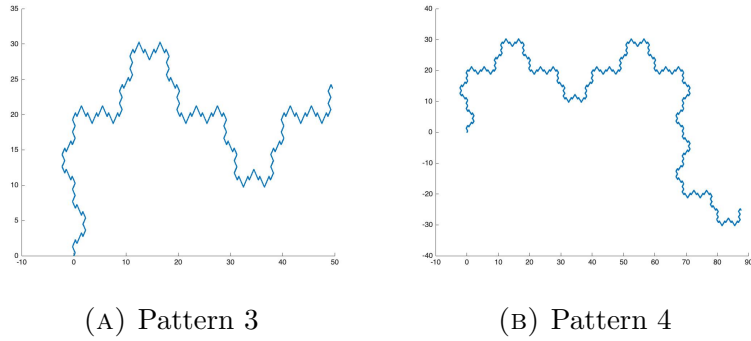


FIGURE 5. Patterns after 150 and 300 collisions

We can see that a pattern emerges after 1000 collisions.

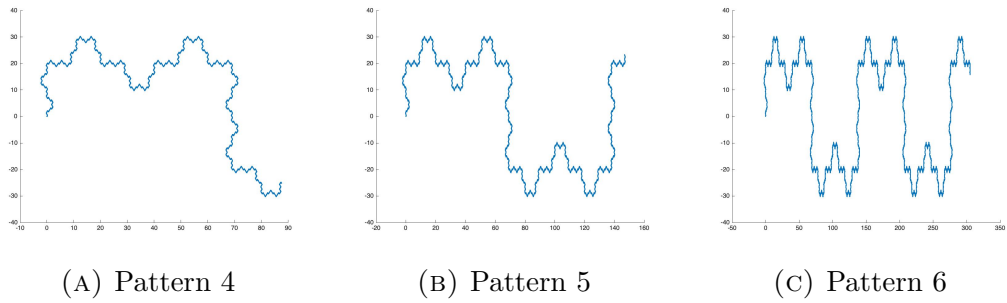
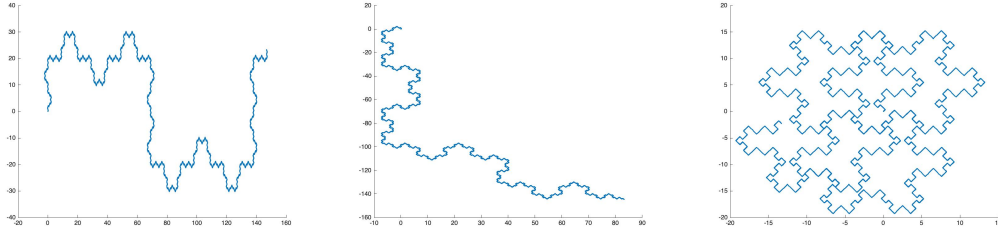


FIGURE 6. Patterns after 300, 500, and 1000 collisions

5. EXPERIMENTS: BOUNCING SQUARE BLOCKS

Experiment 1: We observe at least 3 types of motion.
 Billiard bouncing square blocks with given initial slope for 500 collisions.

- (a). Initial slope = 1.732
- (b). Initial slope = 1.618
- (c). Initial slope = 1.414



(A) Initial slope = 1.732 (B) Initial slope = 1.618 (C) Initial slope = 1.414

FIGURE 7. Patterns after 500 collisions, for 3 types of motion

We notice three types of patterns from the first 500 collisions, as illustrated in figure 7.

- (1). Recurrent (e.g. when initial slope = 1.414)
- (2). Divergent in a quasi-periodic way (e.g. when initial slope = 1.732)
- (3). Divergent rapidly (e.g. when initial slope = 1.618)

In the above classification, we use the term quasi-periodic in (2). We say that the motion is quasi-periodic if there is a T , a quasi-period τ , and a small number $\epsilon > 0$, such that the y -coordinate $y(t)$ of the particle at time t satisfies

$$d(y(t + \tau), y(t)) \leq \epsilon$$

for all $t > T$. The number τ can depend on the time t .

Experiment 2

Billiard bouncing square blocks with given initial slope for 213 collisions.

Initial slope = 1.718

Position at iteration 213 is $x = -0.2500$, $y = 0.4355$

Position at first iteration is $x = 0.2500$, $y = 0.4295$

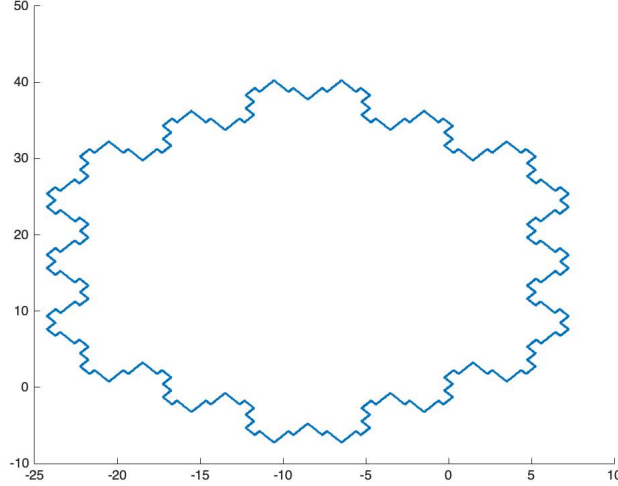


FIGURE 8. Recurrent motion

We observe recurrent motion. Figure 8 shows the particle almost returns to original position.

6. HIDDEN MARKOV PROCESS

Let us review some important definitions and notations in hidden Markov processes. A sequence of discrete random variables $\{C_t: t = 1, 2, 3, \dots\}$ is a Markov chain if it satisfies the Markov property,

$$Pr(C_{t+1} | C_1, C_2, C_3, \dots, C_t) = Pr(C_{t+1} | C_t)$$

Consider a stochastic process $\{X(t): t = 1, 2, 3, \dots\}$ in which the probability distribution of $X(t)$ is determined by the unobserved state C_t of a Markov chain. Let $\underline{C}^{(t)}$ and $\underline{X}^{(t)}$ be the sequence of values from time 1 to time t of C_t and $X(t)$, respectively. Suppose the Markov chain $\{C_t\}$ has m states. Then, the stochastic process $\{X(t): t = 1, 2, 3, \dots\}$ is an m -state Hidden Markov model if for all t ,

$$(1) \quad Pr(X(t) | \underline{X}^{(t-1)}, \underline{C}^{(t)}) = Pr(X(t) | C_t)$$

In applications, the Markov chain $\{C_t\}$ can represent a coin process or some hidden conditions. For example, the daily return of stock prices $\{X(t)\}$, is influenced by the market condition $\{C_t\}$, a Markov chain which is not directly observable. To illustrate the notation, we denote the history of the coin up to time $t = 4$ by $C^{(4)} = \{C_1, C_2, C_3, C_4\}$. The main assumption for Hidden Markov model is

$$Pr(X(4) = 1 | X(1), X(2), X(3), C^{(4)}) = Pr(X(4) = 1 | C_4)$$

Notation: We denote the probability for the coin at time t by

$$\delta_i = \Pr(C_t = i) \quad \text{for } t = 1, 2, 3, \dots, T.$$

Let p_i be the probability mass function of $X(t)$ when the Markov chain is in state i at time t .

$$p_i(x) = \Pr(X(t) = x | C_t = i).$$

As a consequence, if there are m possible outcomes for the coin, then

$$\begin{aligned} \Pr(X(t) = x) &= \sum_{i=1}^m \Pr(X(t) = x | C_t = i) \Pr(C_t = i) \\ &= \sum_{i=1}^m \delta_i p_i(x) \end{aligned}$$

Key assumption: $\{C_t\}$ is a **homogeneous** Markov chain. That means,

$$\Pr(C_{t+1} = j | C_t = i)$$

does not depend on t . It is a number that depends on i and j . We define

$$\Gamma(i, j) = \Pr(C_{t+1} = j | C_t = i).$$

Γ is a probability transition matrix, whose entry in row i and column j is the probability that the coin at time $t + 1$ will be in state j , given that the coin is in state i at time t .

The likelihood at time $t = T$ is given by

$$L_T = \Pr(X(1) = x_1, X(2) = x_2, X(3) = x_3, \dots, X(T) = x_T).$$

The likelihood can be conveniently expressed in matrix notation,

$$(2) \quad L_T = \vec{\delta} \mathbf{P}(x_1) \cdot \mathbf{\Gamma P}(x_2) \cdot \mathbf{\Gamma P}(x_3) \cdot \dots \cdot \mathbf{\Gamma P}(x_T) \vec{1},$$

where $\mathbf{P}(x)$ is defined to be the diagonal matrix whose (i, i) entry is $p_i(x)$. Here, $\vec{1}$ denotes a column vector of ones.

To estimate the parameters of the model, we maximize the likelihood. This can be done by the Baum-Welch algorithm ([8], [9], [11], [14]), which is in effect one of the earliest instances of the EM algorithm ([6]). In the E step of the algorithm, we infer the probabilities of the hidden states C_t from the observations $X(t) = x_t$. Then, in the M step, we update the parameters of the model by maximizing the likelihood, conditional on the hidden states.

The computation in the M step is straightforward. The challenging part is in the E step of the EM algorithm. We include the details in section 8 for completeness.

7. MODEL RESULTS

In this section, we describe how a 3-state hidden Markov process is adequate to model a dynamical property of the periodic wind-tree model.

Let $d(k)$ be the distance of the particle from the origin, after k collisions with the square blocks. Let $m = \tan \theta$ be the initial slope. Let $D(\theta)$ be the minimum distance of the particle from the origin, between the 500^{th} and 1000^{th} collision, when the initial slope is $m = \tan \theta$. To be precise,

$$D(\theta) = \min\{d(k) : 500 \leq k \leq 1000\}.$$

We consider $T = 300$ values of m , starting from $\tan \theta_1 = 1.4140$, goes up by increment of 0.0025, ending at $\tan \theta_T = 2.1615$. Explicitly, the 300 values are:

$$1.4140, 1.4165, 1.4190, 1.4215, 1.4240, \dots, 2.1590, 2.1615.$$

For each value of m , we compute $\log D(\theta)$. The random process $X(t)$ is defined by

$$X(1) = \log D(\theta_1), X(2) = \log D(\theta_2), \dots, X(T) = \log D(\theta_T),$$

where $\theta_1 = 1.4140$, $\theta_2 = 1.4165$, $\theta_3 = 1.4190$, \dots , $\theta_T = 2.1615$.

To model this random process $\{X(t) : 1 \leq t \leq T\}$, we use a hidden Markov model with 3 states, so that conditional on state $C_t = j$, the random variable $X(t)$ is normally distributed with mean μ_j and variance σ_j^2 . The parameters of the 3-state normal-HMM are estimated from the observations $\{X(t), t = 1, 2, 3, \dots, T\}$.

To start the EM algorithm, the probability transition matrix Γ is initialized so that all three diagonal entries are set to 0.8 and all the off-diagonal entries are set to 0.1. This initialization is arbitrary.

By applying the EM algorithm, the parameters of the model after 15 iterations are:

$\mu_1 = -0.613$	$\sigma_1 = 0.13139$
$\mu_2 = 1.9753$	$\sigma_2 = 0.10825$
$\mu_3 = 4.7825$	$\sigma_3 = 1.1217$

with the following 3×3 probability transition matrix Γ ,

$$\begin{array}{ccc} 0 & 1.00 & 0 \\ 0.1262 & 0 & 0.8738 \\ 0 & 1.00 & 0 \end{array}$$

Notice the matrix Γ has an appealing structure. Only two of the nine entries are strictly between 0 and 1.

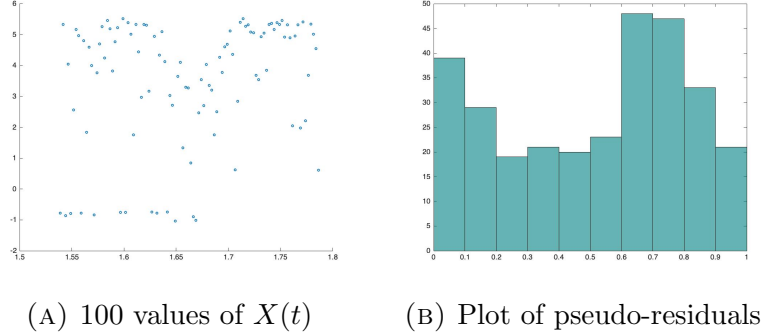


FIGURE 9. The series and the pseudo-residuals

We can interpret the states in the 3-state hidden Markov model so that the states match the 3 types of patterns we observed in Experiment 1 from section 5. For example, state 1 can represent recurrent motion. We caution against the tendency to over-interpret, since we know of no convincing reason to give physical interpretation to each state. Models need not be interpretive; they can be useful as empirical models. We are using model as empirical model, in the sense Cox ([7]) uses that term.

In generalized linear models, it is common to use residuals as a tool to check the validity of the fitted model. The residuals are calculated from the expected observation and the actual observation. Pseudo-residuals are intended to fulfill the same role as the residuals, in the context of HMMs. The uniform pseudo-residual for the observation x_t is

$$u_t = Pr(X(t) \leq x_t) = F_{X(t)}(x_t),$$

where $X(t)$ is a continuous random variable and F is its cumulative distribution function (c.d.f.). The diagrams of the pseudo-residuals and the series $X(t)$ are shown (for ease of viewing, the figure shows 100 values of $X(t)$, from $t = 1.5390$ to 1.7865). We check the histogram for the pseudo-residuals, as a model diagnostic. None of the 10 bins deviate significantly from 30. The highest value among the 10 bins is 48, and the smallest is 19. We conclude that the model is adequate.

8. PROBABILITIES OF THE HIDDEN STATES

In this section, we state and prove the main proposition for computing the probabilities for the hidden states C_t in the E step of the EM algorithm. We include the details here for completeness. Before we can state the proposition, we need to describe the forward and backward probabilities. For the sake of clear exposition, we illustrate

this with 5 time periods. Suppose $T = 5$. The likelihood at time $t = T$ is given by

$$\begin{aligned} L_T &= Pr(X(1) = x_1, X(2) = x_2, X(3) = x_3, X(4) = x_4, X(5) = x_T) \\ &= \vec{\delta} \mathbf{P}(x_1) \cdot \mathbf{\Gamma P}(x_2) \cdot \mathbf{\Gamma P}(x_3) \cdot \mathbf{\Gamma P}(x_4) \cdot \mathbf{\Gamma P}(x_5) \vec{1}, \end{aligned}$$

where we applied formula (2) from section 6 in the last line.

Forward probabilities

$$\begin{aligned} \vec{\alpha}_1 &= \vec{\delta} \mathbf{P}(x_1), & \vec{\alpha}_2 &= \vec{\alpha}_1 \mathbf{\Gamma P}(x_2) \\ \vec{\alpha}_3 &= \vec{\alpha}_2 \mathbf{\Gamma P}(x_3) \\ \vec{\alpha}_4 &= \vec{\alpha}_3 \mathbf{\Gamma P}(x_4) \end{aligned}$$

Backward probabilities

$$\begin{aligned} \vec{\beta}_5 &= \vec{1}, & \vec{\beta}_4^T &= \mathbf{\Gamma P}(x_5) \vec{1}^T \\ \vec{\beta}_3^T &= \mathbf{\Gamma P}(x_4) \mathbf{\Gamma P}(x_5) \vec{1}^T \\ \vec{\beta}_2^T &= \mathbf{\Gamma P}(x_3) \mathbf{\Gamma P}(x_4) \mathbf{\Gamma P}(x_5) \vec{1}^T \\ \vec{\beta}_1^T &= \mathbf{\Gamma P}(x_2) \mathbf{\Gamma P}(x_3) \mathbf{\Gamma P}(x_4) \mathbf{\Gamma P}(x_5) \vec{1}^T \end{aligned}$$

The forward and backward probabilities are related to the likelihood.

$$\vec{\alpha}_t \vec{\beta}_t^T = L_T = Pr(X(1) = x_1, X(2) = x_2, \dots, X(T) = x_T)$$

Properties of forward and backward probabilities

The j -th entry of $\vec{\alpha}_t$ is

$$\alpha_t(j) = Pr(X(1) = x_1, X(2) = x_2, X(3) = x_3, \dots, X(t) = x_t, C_t = j)$$

The i -th entry of $\vec{\beta}_t$ tells us that

$$\beta_t(i) = Pr(X(t+1) = x_{t+1}, \dots, X(T) = x_T, | C_t = i)$$

One property of the HMM is the conditional independence of $X(t+1)$ and $\{X(t+2), X(t+3), \dots, X(T)\}$ given C_{t+1}

$$\begin{aligned} &Pr(X(t+1), X(t+2), X(t+3), \dots, X(T) | C_{t+1}) \\ &= Pr(X(t+1) | C_{t+1}) Pr(X(t+2), X(t+3), \dots, X(T) | C_{t+1}) \end{aligned}$$

We will use these two properties of HMM:

For any integer $T \geq t + 1$,

$$\begin{aligned} & Pr(X(1), X(2), \dots, X(T), C_t, C_{t+1}) \\ &= Pr(X(1), X(2), \dots, X(t), C_t) Pr(C_{t+1} | C_t) \\ &\quad \cdot Pr(X(t+1), \dots, X(T) | C_{t+1}) \end{aligned} \tag{P1}$$

and

$$\begin{aligned} & Pr(X(t+1) | C_{t+1}) Pr(X(t+2), \dots, X(T) | C_{t+1}) \\ &= Pr(X(t+1), \dots, X(T) | C_{t+1}) \end{aligned} \tag{P2}$$

Proposition 5

$$\begin{aligned} & Pr(C_{t-1} = j, C_t = k | X(1), X(2), X(3), \dots, X(T)) \\ &= \frac{\alpha_{t-1}(j) \Gamma(j, k) p_k(x_t) \beta_t(k)}{L_T} \end{aligned}$$

Proof of the Proposition:

$$\begin{aligned} & Pr(C_{t-1} = j, C_t = k | X(1), X(2), X(3), \dots, X(T)) \\ &= \frac{Pr(X(1), \dots, X(T), C_{t-1} = j, C_t = k)}{L_T} \\ &= Pr(X(1), \dots, X(t-1), C_{t-1} = j) Pr(C_t = k | C_{t-1} = j) \\ &\quad \cdot Pr(X(t), \dots, X(T) | C_t = k) \cdot \frac{1}{L_T}, \tag{by (P1)} \\ &= \alpha_{t-1}(j) \Gamma(j, k) \cdot Pr(X(t), \dots, X(T) | C_t = k) \cdot \frac{1}{L_T} \\ &= \alpha_{t-1}(j) \Gamma(j, k) \left(Pr(X(t) | C_t = k) Pr(X(t+1), \dots, X(T) | C_t = k) \right) \cdot \frac{1}{L_T} \tag{by (P2)} \\ &= \alpha_{t-1}(j) \Gamma(j, k) p_k(x_t) \beta_t(k) \cdot \frac{1}{L_T} \end{aligned}$$

End of Proof.

9. Acknowledgements

The first author is grateful for the financial support of an Undergraduate Summer Research Program (USRP) grant from DePaul University.

The second author is grateful that the Undergraduate Summer Research Program (USRP) grant from DePaul University makes this research project possible.

REFERENCES

- [1] Vincent Delecroix, and Anton Zorich, *Cries and whispers in wind-tree forests*, What's next?, Ann. of Math. Stud., 205, Princeton Univ. Press, Princeton, NJ, (2020), pp. 83–115.
- [2] Vincent Delecroix, Pascal Hubert, and Samuel Lelièvre, *Diffusion for the periodic wind-tree model*, Ann. Sci. Éc. Norm. Supér. (4), **47** (2014), no. 6, 1085–1110.
- [3] Vincent Delecroix, *Divergent trajectories in the periodic wind-tree model*, J. Mod. Dyn. **7** (2013), no. 1, 1–29.
- [4] Krzysztof Fraczek, and Corinna Ulcigrai *Non-ergodic Z-periodic billiards and infinite translation surfaces*, Invent math (2014) 197:241–298.
- [5] Nikolai Chernov, and Roberto Markarian, *Chaotic billiards. Mathematical Surveys and Monographs*, 127. American Mathematical Society, Providence, RI, (2006). xii+316 pp.
- [6] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, J. Roy. Statist. Soc. Ser. B **39**, (1977), no. 1, 1–38.
- [7] D. R. Cox, *Role of models in statistical analysis*, Statist. Sci. **5**, (1990), no. 2, 169–174.
- [8] L. E. Baum, and J. A. Eagon, *An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology*, Bull. Amer. Math. Soc. **73**, (1967), 360–363.
- [9] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, *A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains*, Ann. Math. Stat. **41**, (1970), 164–171.
- [10] J. Hardy, and J. Weber, *Diffusion in a periodic wind-tree model*, J. Math. Phys. **21**, (1980), no. 7, 1802–1808.
- [11] B.-H. Juang, and L. R. Rabiner, *Hidden Markov models for speech recognition*, Technometrics **33**, (1991), no. 3, 251–272.
- [12] Serge Tabachnikov, *Geometry and billiards*, Student Mathematical Library, 30. American Mathematical Society, Providence, RI, (2005). xii+176 pp.
- [13] Anton Zorich, *Flat surfaces in Frontiers in Number Theory*, Physics and Geometry. Volume 1: On random matrices, zeta functions and Dynamical Systems, Springer (2006), pp. 439–586.
- [14] Walter Zucchini, Iain L. MacDonald, and Roland Langrock, *Hidden Markov models for time series*, Second edition, Monographs on statistics and applied probability, vol 150. Chapman & Hall, London, (2016).

DEPARTMENT OF MATHEMATICAL SCIENCES, DEPAUL UNIVERSITY, CHICAGO, IL 60614, USA

Email address: `eauyeun1@depaul.edu`

Email address: `nick@hmcproducts.com`