Markov Chains and Traffic Analysis

Emanuel Indrei
August 23, 2006

Department of Mathematics, Georgia Institute of Technology

Abstract

In this paper, we use Markov chains to construct a theoretical traffic system. The paper is organized into three parts: The first two deal with the construction of two spaces in which objects may interact. The third part analyzes the evolution of one particular object. Using bounds given by the law of iterated logarithm and the central limit theorem, we prove that after a large number of time steps, the probability of locating this object in the traffic network diminishes to zero. We conclude with several suggestions on the evolution of multiple objects.

1 Introduction

Mathematical models are commonly used to organize a large collection of data and make future predictions about specific or general outcomes of natural phenomena. This paper investigates the traffic phenomenon. The author often wondered whether we could somehow employ tools from probability theory in order to prove results about the evolution of a traffic system. An accurate model would have many beneficial implications. For instance, it may provide answers to some important questions. One of them being: How many lanes should we build when designing highway systems for various cities in order to maximize flow efficiency and minimize congestion? In this paper, we endeavor to construct a model using Markov chains.

In our analysis, a segment from A to B represents a distance that can be further partitioned into a finite number of sub-segments. Hence, when an object is travelling from A to B, we say that the object travels through all of the sub-segments: $\{A = a_1, a_2, ..., a_{i-1}, a_i, a_{i+1}, ..., a_n = B\}$ and we represent the length of this path by the following equation:

$$L = \sum_{i=2}^{n} (a_i - a_{i-1}). \quad (1)$$

In our model, we want to partition a one way highway system into identical sub-segments so that

$$L = n(a_i - a_{i-1}). \quad (2)$$

We will define a random variable that assigns three possible speeds to objects interacting in our simulation. The speeds will be defined in terms of the number of sub-segments travelled over a specified time period. Hence, objects travelling at higher speeds travel over a larger number of sub-segments during the same time interval. We will define the Markov process in terms of a variant of this random variable.

2 Constructing the Real Space $\Re$

Consider a $m \times n$ configuration matrix A. Place k objects in distinct slots of this configuration matrix where $k \in \{\mathbb{N} \bigcup \{0\} : k \leq mn\}$. Define the real space $\Re$ as the set consisting of the set of all distinct configuration matrices with at most k objects. Hence, we formalize this with the following definitions.

Definition 1. *S_k denotes the set of all configuration matrices with exactly k distinct elements.*

Definition 2. $\Re = \{\cup S_k : k \leq mn\}$.

We note that the empty $m \times n$ matrix is also an element of $\Re$. This is precisely the case when $k = 0$. Now we are ready to prove the following result.

Lemma 1. $|\Re| = \displaystyle\sum_{k=0}^{mn} \binom{mn}{k}$.

Proof. The set S_k consists of all distinct $m \times n$ configuration matrices with k objects. Since the number of configuration matrices with k objects is just the number of combinations of size k from a collection of size mn, we have that $|S_k| = \binom{mn}{k}$. Hence, S_k is finite and this is true for all $k \leq mn$. Furthermore, if $r \neq h$, then $S_r \cap S_h = \varnothing$ since two matrices with a different number of entries can never be equal to one another. Now by Definition 2 and from the principle of inclusion/exclusion in set theory we have that $|\Re| = |S_1 \cup S_2 \cup ... \cup S_{mn}| = \displaystyle\sum_{k} |S_k| - \sum_{k<j} |S_k \cap S_i| + \sum_{k<i<j} |S_k \cap S_i \cap S_j| - \sum_{k<i<j<l} |S_k \cap S_i \cap S_j \cap S_l| + ... + (-1)^{n+1} |S_1 \cap ... \cap S_{mn}| = \sum_{k} |S_k| = \sum_{k} \binom{mn}{k}$. □

Making the transition into the $xy - plane$

Sometimes it will be convenient to locate elements in $\Re$ by using standard Cartesian coordinates rather than matrix notation. Since an element in the a_{ij} entry of A is located in the ith row and jth column, we say that this particular element is an object located in the box whose $xy - plane$ coordinates are given by $(i-1, j-1), (i-1, j), (i, j-1), (i, j)$. Let's look at a simple 4×5 matrix B to get a better feel for the $xy - plane$ transition. Let's say that the b_{23} entry is non-empty. This means that we have an object located in the box outlined by the following coordinates in the first quadrant of the $xy - plane$: $(1, 2), (1, 3), (2, 2), (2, 3)$. The reader is encouraged to draw this example and to notice that the rows are being numbered on the positive $x - axis$ and the columns are being numbered on the positive $y - axis$. This transformation is analogous to taking a regular $m \times n$ matrix, alligning the top left edge of this matrix with the origin and the $xy - axes$ and then rotating this matrix counterclockwise by 90 degrees. Hence, the rows represent highway lanes and the columns represent the length of the highway segment. See Figure 1 below.

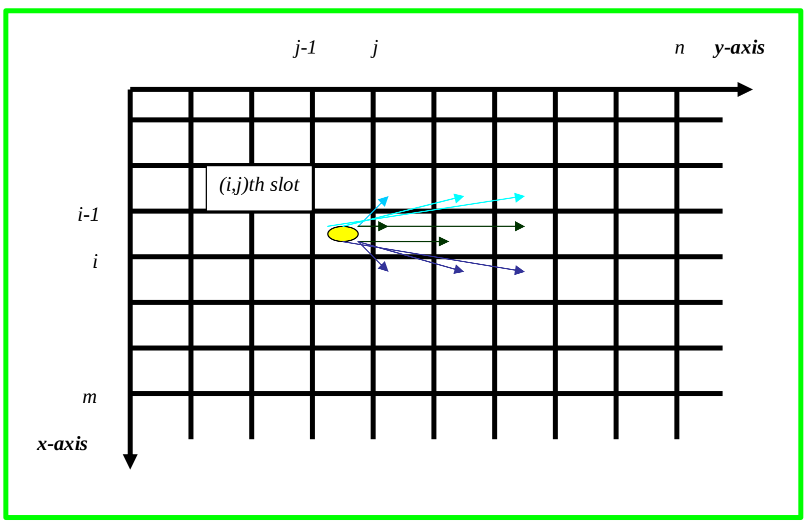

Figure 1.

Rules of Behavior in $\Re$

In real life traffic, objects generally move. We will thus treat our traffic simulation as a dynamic system. An object located at time $t = t_0$ in the $(i, j)th$ slot of matrix A_0 must transition to a new location at time $t = t_1$. Once the object makes a transition, we have a new configuration matrix A_1 that represents the object's location at time $t = t_1$.

The following questions may arise: Where did A_1 come from? What are its dimensions? Well, in order to address the first question, we need to address the second. A_1 represents a configuration matrix; therefore, if we assume that A_0 is an $m \times n$ matrix, then it makes sense to say that A_1 is also an $m \times n$ dimensional matrix. Furthermore, if we assume that A_0 has only one non-empty entry in the $(i, j)th$ slot, then it makes sense to say that A_1 should have at least one non-empty entry. In general, if A_0 has d objects in d slots then A_1 also has d objects in d slots. For simplicity we will assume that $d = 1$. Hence, we have established that A_1 is an element of $\Re$. But now the following question arises: In what slot should this non-empty entry be in A_1?

Before we answer this question, we need to talk a little bit more about real life traffic. When we are driving on a major interstate, almost everyone would agree that we usually drive forward. In our simulation, we ignore accidents, construction zones, road blocks, etc. and focus rather on a smooth flowing traffic network. By smooth, we mean that there are no outside forces interferring with our traffic network. With this in mind it doesn't make sense to say that an object in the $(i, j)th$ slot of A_0 will transition to the $(f(i), j-1)$ slot of A_1 where $f : \{0, 1, 2, ..., m-1, m\} \mapsto \{0, 1, 2, ..., m-1, m\}$ since if we are in a helicopter which is perpendicular to the interstate we generally would observe cars travelling from left to right and our matrix is precisely an aerial (two dimensional) view of the traffic network. Now suppose that we are driving in a 5 lane highway system on the middle lane and we want to take the next exit. (Imagine that we are in a helicopter which is alligned with the highway so that the exit lane is the first lane from the right and the middle lane is the third lane from the right). In order to get to the exit lane we need to first get through the lane between the middle lane and the exit lane (i.e., the second lane from the right). We can't jump over lanes in real life; hence, it makes sense to say that at each time step, we are only allowed to move one lane at a time.

With all of this in mind, we are now ready to mathematically consider lane transitions represented as follows:

$$(i, j) \bowtie (f(i), j + \sigma(object)). \tag{3}$$

This means that an object in the $(i, j)th$ slot in some matrix $A_0 \in \Re$ will end up in the $(f(i), j + \sigma(object))th$ slot of some matrix $A_1 \in \Re$ during one time step $(A_0 \neq A_1)$, where

$$\sigma(object) = \begin{cases} 1, & \text{with probability } p \\ 2, & \text{with probability } q \\ 3, & \text{with probability } r \\ p+q+r=1 & \end{cases} \tag{4}$$

and

$$f : \{i\} \mapsto \{i, i+1, i-1\}. \tag{5}$$

Figure 1 displays all possible transitions for the object in position (i, j). But as we look at the scenario, the following question arises: What happens at the boundaries?

Boundary Conditions

In order to come up with a logical way of determining boundary conditions, we refer to the behavior of real life traffic. The introduction of this paper makes it clear that we are dealing with a one way highway system. So if we imagine a 5 lane highway system (imagine a 5×2000 matrix) in which the cars are travelling from left to right, we immediately see that the cars in the top lane cannot escape the system by transitioning to a lane that is above them. The same rule applies for cars at the very bottom. They can't simply transition into an imaginary lane that is below them. Hence, there are two natural boundaries and so we say that the k_1 cars in the first row of the initial $m \times n$ configuration matrix A and the k_m cars in the mth row of A both adhere to Equation (4), but for the k_1 vehicles, Equation (5) becomes

$$f_1 : \{i\} \mapsto \{i, i-1\}. \tag{6}$$

and analogously for the k_m vehicles, Equation (5) becomes

$$f_m : \{i\} \mapsto \{i, i+1\}. \tag{7}$$

Now we turn to the vertical boundaries (as seen from the matrix). Consider Interstate 85 ($I-85$) in Georgia and the segment of $I-85$ that ranges from Exit 115 to Exit 120. We know that usually there are cars on $I-85$ that are not on this segment since $I-85$ is a rather long highway (664 miles to be exact). Well, we also know that if a car is travelling continuously from Atlanta to Greenville solely on $I-85$, then it has to pass through this segment. With this in mind, we can establish that cars are consistently flowing into our system ($influx$) and cars are consistently flowing out of our system ($outflux$). Moreover, in our simulation we have the following definition.

Definition 3. *influx(System)=outflux(System).*

Let's consider once again a long 5 lane highway system (let's say a 5×5000 matrix this time) in the first quadrant of the $xy - plane$ as outlined on page 3. If we assume that $k = 200$ (i.e., we have 200 cars in this system at time $t = 0$), then as soon as we let this process evolve we keep track of car X_r which is the car furthest to the right in the matrix (i.e., X_r is in the box with the largest $y - coordinate$ on the $xy - plane$). Whenever X_r hits the boundary outlined by $y = 5000$ and $0 \leq x \leq 5$ on the $xy - plane$ (note that $m = 5$ in this case and $n = 5000$), we say X_r flows out of the system. But since $influx(System) = outflux(System)$, there must be a car X_i flowing into the system by crossing the horizontal boundary outlined by $0 \leq x \leq 5$ and $y = 0$. Hence, we see that the number of cars in the system is always a constant. In fact we can say more: the number of cars in the system is exactly equal to k, the initial number of cars in the system. Since the evolution of the traffic system is just a transition between different elements of $\Re$ we have the following result.

Lemma 2. *If at time $t = 0$ there are k objects in the traffic simulation, then the system evolves over the space $S_k \subseteq \Re$. (Recall that Definition 1 states that S_k denotes the set of all configuration matrices with exactly k distinct objects).*

Moreover, we know from the proof of Lemma 1 that

$$|S_k| = \binom{mn}{k}. \tag{8}$$

So this means that once we make our selection of how many cars we want to start out with, we can identify the subset of $\Re$ in which we are working and the cardinality of this subset is significantly less than the cardinality of $\Re$. Hence, we see that $\Re$ has a very nice structure. Moreover, we can define an equivalence relation on $\Re$:

Definition 4. *Given matrices A_1, A_2, $A_3 \in \Re$, let $A_1 \equiv A_2$ if A_1, $A_2 \in S_k \subseteq \Re$.*

Lemma 3. *Definition 4 defines an equivalence relation.*

Proof. 1. $A_1 \equiv A_1$ since A_1 has the same number of objects as A_1.

2. Assume $A_1 \equiv A_2$. This means that A_1 and A_2 are both in the set S_k and so they both have k objects. This implies that $A_2 \equiv A_1$.

3. Assume $A_1 \equiv A_2$ and $A_2 \equiv A_3$. This means that A_1 and A_2 are in some set $S_r \subseteq \Re$ with r objects and that A_2 and A_3 are in some set $S_h \subseteq \Re$ with h objects. In the proof of Lemma 1, we established that if $r \neq h$, then $S_r \cap S_h = \varnothing$. Therefore, by the contrapositive we have that $r = h$ since in our case $A_2 \in S_r$ and $A_2 \in S_h$. This implies that $A_1 \equiv A_3$. □

Now that we have defined an equivalence relation on $\Re$ we are guaranteed equivalence classes which provide us with a decomposition of $\Re$ as a union of mutually disjoint subsets. It is readily visible that the decomposition consists of all the unique sets which contain matrices with the same number of non-empty entries. Of course in our case, these non-empty entries correspond to objects, which in turn correspond to cars. Now that we have investigated the structure of $\Re$, we are ready to move onto the construction of the camera space which will be a crucial tool for proving various results about the evolution of this system.

3 Constructing the Camera Space Ψ

Choose positive integers m and n such that they outline the desired dimensions of a highway segment which is represented as an $m \times n$ configuration matrix $A \in \Re$. Next select a positive integer v such that $v \ll n$. Now consider the $m \times v$ **empty** configuration matrix C. We represent C on the $xy-plane$ through the same method which was used to represent elements of $\Re$ on the $xy-plane$ outlined in Section 2.

Now we construct Ψ from $\Re$ through the following procedure:

Step 1. Select a configuration matrix $T \in \Re$ such that it has exactly k elements and represent it by the method outlined in Section 2 on the $xy-plane$.
Step 2. Select an **empty** configuration matrix C and place it in its respective coordinates on the $xy-plane$ as outlined in the preceding paragraph. Note that all entries of C overlap a subset of the entries of T. Some entries of T may have objects in them whereas others may not. Nevertheless, there are exactly k objects in the entries of T.
Step 3. If T has an object in its $(i,j)th$ entry and if it overlaps the $(i,j)th$ entry of C, then let C absorb an exact replica of this object in the entry of intersection.
Step 4. Repeat Step 3 with all other distinct $mn-1$ entries of T. (Recall that T is an $m \times n$ matrix and so we have to check a total of mn entries.)
Step 5. After carrying out Step 4, denote the resulting configuration matrix C by C'. Note that C' may be **nonempty** whereas C is **always** empty.

Definition 5. *The camera space Ψ is the space consisting of the set of all possible configurations for C'.*

It is clear that all possible configurations of C' have at most k objects, where k denotes the number of objects initially selected for matrices in $\Re$. Hence, we have that

$$|\Psi| = \sum_{k=0}^{mv} \binom{mv}{k}. \tag{9}$$

Rules of Behavior and Boundary Conditions in Ψ

The same exact rules we developed in Section 2 for objects interacting in $\Re$ also apply for objects interacting in Ψ, with one exception. Recall Equation (3) in Section 2 defined lane transitions as

$$(i, j) \bowtie (f(i), j + \sigma(object)). \tag{10}$$

We redefine lane transitions for objects interacting in Ψ as

$$(i, j) \bowtie (f(i), j + \delta(object)). \tag{11}$$

Recall what this means. Namely, that objects in the $(i, j)th$ slot of some matrix transition to the $(f(i), j + \delta(object))$ slot of some other matrix. Note that the only difference between Equation (10) and (11) is the symbol $\delta(object)$. We define $\delta(object)$ as the following.

$$\delta(object) = \begin{cases} -1, & \text{with probability } p \\ 0, & \text{with probability } q \\ 1, & \text{with probability } r \\ p + q + r = 1 & \end{cases} \tag{12}$$

$\delta(object)$ is a normalized version of $\sigma(object)$. Recall that $\sigma(object)$ assigns objects interacting in $\Re$ a certain speed at each time step. Likewise, we say that $\delta(object)$ assigns a certain speed to objects interacting in Ψ. (From here on, we denote objects interacting in $\Re$ or Ψ simply as objects in $\Re$ or Ψ). Objects in $\Re$ can only disappear completely from the traffic system when they cross the line $y = n$ for $0 \leq x \leq m$, whereas objects in Ψ can disappear temporarily from the very start and then show up again after a couple of time steps. This can happen because of the negative term in $\delta(object)$.

Visualizing Ψ

Imagine a real traffic system. Locate one car at time $t = 0$ and keep track of this car for a fixed distance (call it Supcar to distinguish it from the other vehicles). Supcar may speed up, slow down, or remain at a constant speed throughout the expedition. Now think of an Apache helicopter hovering right above the traffic system at time $t = 0$. On the helicopter, there is a camera which has a limited viewing frame so that when it is pointed down towards the interstate, it can only capture a small portion of it. Now imagine that this camera captures Supcar at time $t = 0$. Like Supcar, the helicopter may speed up, slow down, or remain at a constant speed hence its behavior may vary and so there may be some instances in which the helicopter's camera catches Supcar and likewise there may be some instances in which Supcar will be to the left or to the right of the camera on the fixed interval under consideration. In this illustration, the interstate represents $\Re$, Supcar represents some object interacting in $\Re \cup \Psi$, and the camera on the helicopter represents Ψ.

The Disappearance Dilemma

Suppose that $object_1$ travels for a while by transitioning through elements of $\Re$. Then $object_1$ eventually crosses the boundary set up for elements of $\Re$ (recall that this boundary is the line $y = n$ for $0 \leq x \leq m$) and disappears permanently. If we have a matrix $C \in \Psi$ and suppose that the replica of $object_1$ interacting in Ψ happens to be in C right before $object_1$ dissapears permanently out of $\Re$, then we say that the replica of $object_1$ also dissapears out of Ψ.

In the construction of Ψ we had that $v \ll n$. Suppose on the other hand that $v = n$, and suppose that $p = 0$, $q = 1$, and $r = 0$ and $f(i) = i$. Then, if the object is initially located in the $(i, j)th$ slot of some matrix $A \in \Re$ and its duplicate in some matrix $C \in \Psi$, then the duplicate will stay in that slot througout the original's entire expedition and it will eventually disappear out of that slot precisely when the original object reaches the boundary condition outlined for elements of $\Re$.

4 The Evolution of Objects interacting in $\Re$ and Ψ

Markov Chains

A stochastic process denoted as $\{Y_t, t \in \mathbb{N}\}$ is defined as a collection of random variables. In our system, Y_t represents car X's **column** position at time t in the elements of $\Re$. Hence, we have that

$$Y_t = \sum_{i=1}^{t} \sigma_i. \tag{13}$$

Lemma 4. *If* $X_t = \sum_{i=1}^{t} \delta_i$ *is known, then* Y_t *is known for each* t.

Proof.

$$\begin{aligned} Y_t &= \sum_{i=1}^{t} \sigma_i \\ &= \sum_{i=1}^{t} (\delta_i + 2) \\ &= \sum_{i=1}^{t} \delta_i + \sum_{i=1}^{t} 2 \\ &= X_t + 2t \end{aligned} \tag{14}$$

□

We note that the random variable X_t is a sum of random variables of the form δ_i introduced in Section 3. Furthermore, we note that the position of X_{t+1} is only dependent upon X_t with the probabilities p, q, r, where $p+q+r=1$. Therefore, since the conditional distribution of any future state X_{t+1} given the past states X_0, X_1, ... , X_{t-1} and the present state X_t is independent of the past states and depends only on the present state, we have that this stochastic process is indeed a Markov chain.

Random Walks

We have previously shown how traffic may be modeled using Markov chains. Now we consider random walks. A Markov chain is said to be a random walk if for some real number $0 < p < 1$ we have that $P_{i,i+1} = p = 1 - P_{i,i-1}$ where $i \in \mathbb{Z}$. This means that an object in state i transitions to state $i+1$ with a fixed probability p and the same object in state i transitions to state $i-1$ with a fixed probability $1-p$. In our simulation we have that the state of the object at time t corresponds to the matrix $A_t \in \Re$ in which the object is located. But we can't simply identify A_t out of $\Re$ since this is a stochastic process (i.e., the location of the object is determined by probabilities). The best we can do is calculate Y_t. However, in order to define a random walk for our object, we will be working in Ψ since X_t will yield full information about Y_t.

Traffic Application: The Symmetric Case

With this in mind, we infer that in our traffic simulation, car X undergoes a one-dimensional random walk on the integers if $q = 0$, since then it steps either to the right with probability r or to the left with probability $p = 1 - r$. We are interested in the evolution of X after a large number of time steps. If $p > r$ then it is easy to see that X will indeed diverge to the left out of Ψ. Similarly, if $p < r$ then X will diverge to the right out of Ψ. The interesting question is: What happens when $p = r$?

This scenerio is called the symmetric one-dimensional random walk. In this scenerio, one would expect Ψ to keep track of X, since both Ψ and X travel at medium speed on average. However, this is not the case as we will soon see. But before we prove this result, we need some preliminaries.

Lemma 5. $\mathbb{E}(X_t) = 0$ *in the symmetric case where* X_t *is given by Lemma* 4.

Proof.

$$\begin{aligned}\mathbb{E}(X_t) &= \mathbb{E}(\sum_{i=0}^{t} \delta_i) \\ &= \sum_{i=0}^{t} \mathbb{E}(\delta_i) \\ &= 0\end{aligned} \tag{15}$$

since

$$\begin{aligned}\mathbb{E}(\delta_i) &= (-1)(p) + (0)(q) + (1)(r) \\ &= -p + r \\ &= 0\end{aligned} \tag{16}$$

because $p{=}r$. □

This result confirms our previous claim that on average, car X travels at medium speed. In order to go a step further, we need to make use of two well-established results in probability theory.

The Central Limit Theorem

One of the most important results in probability theory is the central limit theorem. This theorem is useful for approximating probabilities for sums of independent random variables. It can be found in many elementary probability textbooks [1].

Theorem 1. *Let X_1, X_2, ... be a sequence of independent, identically distributed random variables, each with mean μ and variance σ^2. Then the distribution of*

$$\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution as $n \to \infty$. That is,

$$\mathbb{P}\{\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} \leq a\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{\frac{-x^2}{2}} dx$$

Law of Iterated Logarithm

The law of iterated logarithm comes in many forms. It tells us that for a one dimensional symmetric random walk, the oscillatory behavior of a random variable tends to increase with time. The original statement is due to A.Ya. Khinchin [2]. Another statement was given by A.N. Kolmogorov [3]. The following is a simpler form of this result (Theorem 3.52 in Breiman [4]):

Theorem 2. $\limsup_{n\to\infty} \frac{|S_n|}{\sigma\sqrt{2nloglog(n)}} = 1$ *(almost surely) where* S_n *is the sum of* n *independent, identically distributed variables with mean zero and finite variance* σ^2.

Dude, where is my car?

Let's imagine the standard first and fourth quadrants of the $xy-plane$. We place a car at the origin at time $t = 0$. Then as the process starts, we let this car undergo a symmetric random walk on the integers by making it equally likely that the car will be one unit up or one unit down at each time step. According to the law of iterated logarithm, we have bounds for this car's random oscillatory behavior. Now assume that the lines $y = a$ and $y = b$ represent the camera space Ψ. This means that whenever the car is between those lines, then it is transitioning through elements of Ψ and $\Re$, but whenever the car is not between those lines, then it is only transitioning through elements of $\Re$. As we have previously stated, we would expect Ψ to keep track of this object, but this is not the case. We state this result as Corollary 1.

Corollary 1. $\mathbb{P}(a \leq X_t \leq b) = 0$ *for large* t*, where*

$$X_t = \sum_{i=1}^{t} \delta_i$$

and

$$\sigma_i = \left\{ \begin{array}{ll} -1, & \textit{with probability } p \\ 0, & \textit{with probability } q \\ 1, & \textit{with probability } r \end{array} \right\}.$$

Proof. Normalizing the equation on the left, we get

$$\mathbb{P}(a \leq X_t \leq b) = \mathbb{P}\left(\frac{a - \mathbb{E}(X_t)}{\sqrt{\mathbb{V}(X_t)}} \leq \frac{X_t - \mathbb{E}(X_t)}{\sqrt{\mathbb{V}(X_t)}} \leq \frac{b - \mathbb{E}(X_t)}{\sqrt{\mathbb{V}(X_t)}}\right). \tag{17}$$

Lemma 5 tells us that $\mathbb{E}(X_t) = 0$. Thus,

$$\begin{aligned}
\mathbb{V}(X_t) &= \mathbb{E}(X_t^2) - (\mathbb{E}(X_t))^2 \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{t} \delta_i\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{t} \delta_i^2\right] + \mathbb{E}\left[\sum_{i\neq j}^{t} \delta_i\delta_j\right] \\
&= \sum_{i=1}^{t} \mathbb{E}(\delta_i^2) + \sum_{i\neq j}^{t} \mathbb{E}(\delta_i\delta_j)
\end{aligned}$$

where $\mathbb{E}(\delta_i\delta_j) = \mathbb{E}(\delta_i)\mathbb{E}(\delta_j)$ since δ_i and δ_j are independent. But in the proof of Lemma 5 we have established that $\mathbb{E}(\delta_i) = 0$. Hence, $\mathbb{E}(\delta_i\delta_j) = 0$. However, $\mathbb{E}(\delta_i^2) = p + r$ since δ_i is not independent of itself. Hence,

$$\mathbb{V}(X_t) = t(p + r)$$

Now for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $t \geq N$,

$$\frac{|a|}{\sqrt{t(p+r)}} < \epsilon$$

and

$$\frac{|b|}{\sqrt{t(p+r)}} < \epsilon.$$

Hence,

$$\mathbb{P}\left(\frac{a}{\sqrt{t(p+r)}} \leq \frac{X_t}{\sqrt{t(p+r)}} \leq \frac{b}{\sqrt{t(p+r)}}\right) \leq \mathbb{P}\,(\,-\epsilon \leq \frac{X_t}{\sqrt{t(p+r)}} \leq \epsilon\,).$$

By the Central Limit Theorem, we have that

$$\lim_{t\to\infty} \mathbb{P}\left(-\epsilon \leq \frac{X_t}{\sqrt{t(p+r)}} \leq \epsilon\right) = \frac{1}{\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} e^{\frac{-x^2}{2}}\, dx.$$

Since, we can do this for any $\epsilon > 0$,

$$\lim_{\epsilon\to 0} \frac{1}{\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} e^{\frac{-x^2}{2}}\, dx = 0.$$

But this implies that

$$\mathbb{P}(a \leq X_t \leq b) = 0.$$

□

A Clarification of an argument used in Corollary 1

In Corollary 1 we let $t \to \infty$ in order to use the central limit theorem. We have to figure out what this means in the context of $\Re$ since $\Re$ is a finite dimensional space whose cardinality we computed in Lemma 1. Well, we said before that once an object disappears out of $\Re$ then it is gone forever and a new object takes its place. But then no object can undergo a symmetric random walk on the integers since after some finite time it will reach the boundary and disappear out of both $\Re$ and Ψ. The way we get around it for the purpose of the argument in Corollary 1 is by saying that once an object hits the boundary it appears again as the new object which is supposed to enter the system in its place (recall influx(system)=outflux(system)). This scenerio is consistent with how we previously defined boundary conditions and we are going to treat the evolution of the system similarly whenever we will let the number of transitions approach ∞. One way to picture this scenerio is by stacking matrices out of $\Re$ on top of each other whenever the object reaches the boundary so that the object will always be interacting in some matrix which belongs to $\Re$. This is analogous to extending a highway system by paving new roads instantaneously whenever a car reaches a point where the pavement terminates (maybe in a desert?).

What does Corollary 1 tell us?

In short, Corollary 1 tells us that after a large number of steps, the oscillations of the car undergoing the symmetric random walk get so large that the probability of locating this car in some element of $\Re$ is zero. Hence, we lose all information with regard to this object's location. This is a surprising result since we expected the object, which on average travels with medium speed, to mostly hang out in the camera space which also travels on average with medium speed. This is not so.

Concluding Remarks

Our analysis this far has only been valid for one object, but we can let two objects undergo one-dimensional symmetric random walks simultaneously. In order to do that, we consider the difference between two random variables X and Y as another random variable Z. The number of collisions is the number of times Z crosses zero. Since we can calculate Z's distribution from the distribution of X and Y, we can think of this as a distribution of recurrence times to zero. The number of collisions in n steps should average out to $\frac{n}{expected\ value\ of\ first\ return}$, i.e., the number of collisions should grow linearly with n with slope $= \frac{1}{expected\ value\ of\ first\ return}$.

For more than 2 cars, we can form the pairwise differences of their positions and then look for the first time any of them is zero. In other words, we are looking at a random variable which is defined as the minimum value of all the pairwise differences. This gets into joint distributions of random variables.

References

[1] Sheldon M. Ross. *Introduction to Probability Models.* Eighth Edition. 2003: Academic Press.

[2] A. Khintchine. *Über einen Satz die Wahrscheinlichkeitsrechnung.* Fundamenta Mathematica, 6:9-20, 1924.

[3] A. Kolmogoroff. *Über das Gesetz des iterierten Logarithmus.* Mathematische Annalen, 101:126-135, 1929.

[4] Leo Breiman. *Probability.* Addison-Wesley, 1968; reprinted by Society for Industrial and Applied Mathematics, 1992. (See Sections 3.9, 12.9 and 12.10)