RANDOM WALKS AND THE PROBABILITY OF RETURNING HOME

ELIZABETH G. OMBRELLARO

Abstract. This paper is expository in nature. It intuitively explains, using a geometrical and measure theory perspective, why a random walker will only pass through the origin infinitely many times in the first and second dimension on $\mathbb{Z}$. The paper also explores why in higher dimensions, the random walker only passes through the origin finitely many times.

Contents

1. Introduction 1
2. Relevant Definitions 2
3. Random Walk 5
Acknowledgments 13
4. Bibliography 13
References 13

1. Introduction

Suppose you play a game where at the flip of a coin you move forwards or backwards one step, depending on whether the coin reads heads or tails. The question then arises, if you were to play this little game, flipping your coin infinitely many times, how many times would you return to the starting point, or the origin.

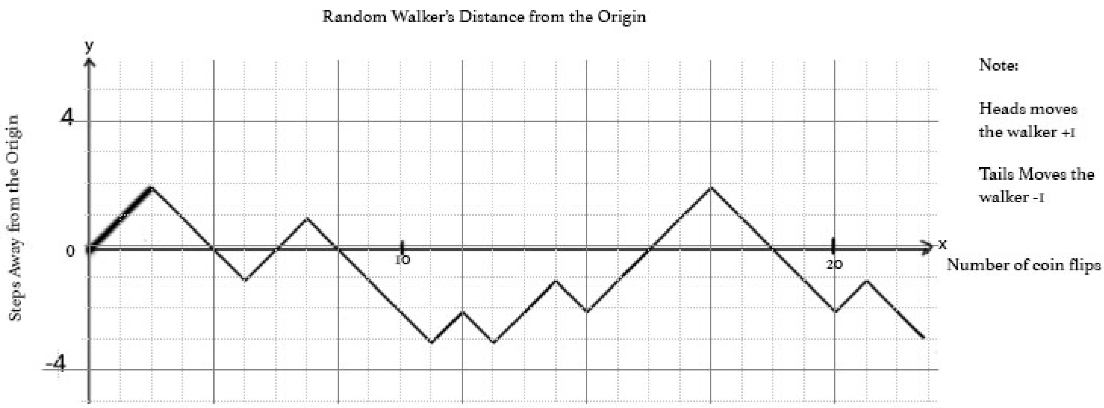
Intuitively, since you have equal probability of getting heads and tails, that means it is as likely to move forwards as backwards, and in a large sample space the amount of heads and tails should be the same, which means that overall you should end back at your starting point or origin infinitely many times.

Now let's say you were not only able to levitate, but also able to go through the ground, or walk along any diagonal path originating from your starting point (neglecting barriers such as the physical ground), in a way that your movement could be represented by a 3-dimensional axis.

With the increase in dimensions, the movement becomes more complicated as with each step you have a new option in each direction of the three dimensions. This in turn raises the question if the walker can end up at the origin infinitely many times. Since the options are numerous it would make sense that the number of returns to the origin would be limited. The goal of this paper is to illustrate how the multi-dimensional random walk varies from the simple case.

Since several natural phenomena, such as protein structures, can be modeled as random walks, an understanding of random walks aides our understanding of complex natural structures too.[5]

Date: November 22, 2018.



2. Relevant Definitions

Definitions 2.1- 2.14 follow from Probability with Martingales [1], whereas Definitions 2.15-2.17 follow from Random Walk and the Discrete Heat Equation [2].

In this section we will discuss relevant definitions of measure theory and probability which will serve us later in our proofs.

These definitions below (2.1-2.13) will give us the groundwork to understand the definition of a Random Variable and an Independent Random Variable, and how we are able to assign values to the likelihood that a Head or Tail is to occur in a coin toss. Definitions 2.15- 2.17 are terms that will be used in the formal proof of the paper.

Definition 2.1 (σ-algebra)**.** A *σ-algebra* is a collection called Σ which is composed of subsets of a set S such that:

(1) if $G \in \Sigma$, then also $G^c \in \Sigma$, therefore Σ is closed under complement.
(2) $S \in \Sigma$, and also $S^c = \emptyset \in \Sigma$
(3) if $G_n \in \Sigma$, then $\bigcup_{n\in\mathbb{N}} G_n \in \Sigma$, therefore it is closed under countable unions.

Remark 2.2. Since Σ is closed under the complement and finite unions, if $G_n \in \Sigma$ then:

$$\bigcap_{n\in\mathbb{N}} G_n = \left(\bigcup_{n\in\mathbb{N}} G_n^c\right)^c \in \Sigma.$$

Definition 2.3 (Borel σ-algebras)**.** Let S be a topological space, the *Borel σ-algebra*, $B(S)$, is the smallest *σ-algebra* taken on the open subsets of S. Define $B = B(\mathbb{R})$.

Example 2.4. Take the topological space $S = \mathbb{R}$, by Definition 2.1.2 $\mathbb{R} \in B(\mathbb{R})$. The examples below are also in $B(\mathbb{R})$.

(1) Singleton: $\{x\}$, $x \in \mathbb{R}$. In the case where $\{x\} = \{1\}$, by definition the open set $(1-\frac{1}{n}\ ,\ 1+\frac{1}{n})$ is in the *Borel σ-algebra*. Therefore $\bigcap_{n\in N}(1-\frac{1}{n}\ ,\ 1+\frac{1}{n}) =$

{1} is also in the *Borel σ-algebra*. According to the same line of reasoning, $\bigcap_{n \in N} (x - \frac{1}{n}\ ,\ x + \frac{1}{n}) = \{x\}$, therefore $\{x\}$ is in the *Borel σ-algebra*.

(2) $\mathbb{Q}$: Since $\mathbb{Q}$ is countable, $\mathbb{Q}$ can be formed from a countable union of singletons. According to Definition 2.1.3, $\mathbb{Q}$ is in the *Borel σ-algebra*. [3]
(3) $\mathbb{R} \setminus \mathbb{Q}$: Since $\mathbb{Q}$ is in the *Borel σ-algebra*, according to Definition 2.1.1, the complement must also be in the *Borel σ-algebra*.

Definition 2.5 (Measurable space)**.** A *measurable space* is a pair (S, Σ), where S is a set, and Σ is a *σ-algebra* on S.

Definition 2.6 (Measure space)**.** Let (S, Σ) be a *measurable space*. Let G_n be a sequence of disjoint sets in Σ, therefore $G_1 \bigcap G_2 ... \bigcap G_n = \emptyset$, and G is equal to the countable union of G_n, where $G \in \Sigma$.

A map $\mu : \Sigma \to [0, \infty]$ is called a *measure* on (S, Σ) if:

(1) $\mu(\emptyset) = 0$.
(2) $\mu(G) = \sum_n \mu(G_n)$, which means μ is countably additive.

If all of these conditions are met, then (S, Σ, μ) is called a *measure space*.

An intuitive example is suppose that each set represents a wall that we paint, and since the set is disjoint, the walls are neither the same nor connected. When we apply the function μ, we are figuring out the painted surface area. When we have no walls, we have no surface to paint, but suppose we want to know the total area painted, we would sum the painted area of each individual wall.

Definition 2.7 (σ-algebra measurable function, mΣ)**.** Let (S, Σ) be a *measurable space*.

A map $h : S \to \mathbb{R}$, where for $A \subseteq \mathbb{R}$, $h^{-1}(A)$ is defined as $\{s \in S : h(s) \in A\}$, is called a σ- algebra measurable function (mΣ) if:

$$h^{-1} : B \to \Sigma$$

This implies that $h^{-1}(A) \in \Sigma$, for every $A \in B$.

Definition 2.8 (Probability measure)**.** Our *measure* μ is called a *probability measure* if $\mu(S) = 1$. We use $\mathbb{P}$ to express a *probability measure*.

Definition 2.9 (Sample Space and Point)**.** In probability theory a *sample space* is a topological space, defined as a set Ω, and a *sample point* is a point ω of Ω.

Definition 2.10 (Family of Events and Events)**.** Define F to be the *σ-algebra* of Ω (It follows the conditions of Definition 2.1). F is defined as the *family of events*, and an *event* is an element of F. In other words an *event* is a *F-measurable* subset of Ω.

In order to have an *experiment* we apply a *probability measure* to our *event* and our *family of events*, therefore we use the probability triple ($\mathbb{P}$, Ω, F).

Intuitively all of this means that Ω is all of the outcomes possible, ω is one specific outcome that occurs, and when we apply $\mathbb{P}(F)$ we find out the probability that an *event* occurs.

Example 2.11. Suppose we have the *Experiment* of flipping a coin twice, the *family of events* is F, and an *event* is an *F-measurable* subset of Ω. In this experiment

$\Omega = \{HH, TT, HT, TH\}$ and

$F = \{\emptyset, \{HH\}, \{TT\}, \{HT\}, \{TH\}, \{HH, HT\}, \{HH, TH\}, \{HH, HT, TT\}, \{HH, TH, TT\}...\}$.

Then suppose we want the *event* of getting at least one head, therefore there exists an $f \in F$, where $f = \{HT, TH, HH\}$. The probability that the *event* f occurs, $\mathbb{P}(f) = 3/4$.

Definition 2.12 (Random variable)**.** If Ω is a *sample space*, F is a *family of events*, and B represents a *Borel σ-algebra*. Then a *random variable* is an element of mF, where mF implies that $X : \Omega \to \mathbb{R}$, and $X^{-1} : B \to F$.

Definition 2.13 (Independent σ-algebras)**.** Sub *σ-algebras* $G_1, G_2, ...$ of F are *independent* if whenever $g_i \in G_i$ with $i \in \mathbb{N}$ and $i_1, ...i_n$ are distinct, then:

$$P(g_{i_1} \bigcap g_{i_2} ... \bigcap g_{i_n}) = \prod_{k=1}^{n} P(g_{i_k}).$$

Definition 2.14 (Independent Random Variables)**.** A *Random Variable* is independent if the *σ-algebras* of the *Random Variable* are independent also. So $\sigma(X_i)$ with $i \in \mathbb{N}$ are independent. Therefore X_i is not influenced by X_{i-1}.

Definition 2.15 (S_n)**.** S_n is the distance from the origin composed of the sum of random variables X_i, where each X_i is not influenced by X_{i-1}. The subscript n denotes the number of steps taken by the random walker.

In this specific case of random walk $\mathbb{P}\{X_i = \pm 1\} = 1/2$, as $X : \{H\} \to 1$ and $X : \{T\} \to -1$. Intuitively this means when we flip a heads we move forward one step and when we flip a tails we move back one step, and it is equally likely both instances will happen.

Definition 2.16 (Expectation)**.** The expected value can be expressed as an average. Intuitively this means adding all the possible values and dividing it by the total number of options. Lawler defines it to be:

$$E[X] = \sum_{z} z * \mathbb{P}\{X = z\}.$$

Looking at examples will clarify as to why it is defined as such.

Example 2.17. (1) In the case of a coin where we are asking the expected value after one toss, we would get:

$$E[S_1] = E[0 + X_1].$$

Remark 2.18. Lawler mentions that E[X+Y]= E[X] + E[Y].

In general when we take a look at the expected average distance, we get:

$$E[Sn] = E[0] + E[X_1] + E[X_2] + ... + E[X_n],$$

Where:

$$E[X_1] = 1 * \mathbb{P}\{X = 1\} + -1 * \mathbb{P}\{X = -1\} = 1/2 - 1/2 = 0.$$

Since all X_i have the same probability of going forward and back, we can replace $E[X_1]$ with $E[X_i]$ for i=1,2,3...

Therefore:

$$E[Sn] = \sum_{i=1}^{n} E[X_i] = 0.$$

(2) In the case with a fair dice, the average value or the expected value we would get is:

$$1 * 1/6 + 2 * 1/6 + ... + 6 * 1/6 = \sum_{k=1}^{6} k/6 = 3.5.$$

(3) In the case where a die is not fair, but weighted such that:

$$\mathbb{P}\{X = 1 \; or \; 2\} = 4/12 \; and \; \mathbb{P}\{X = 3, 4, 5 \; or \; 6\} = 1/12,$$

where X is the value of the roll, then:

$$E[X] = 1 * 4/12 + 2 * 4/12 + 3 * 1/12 + ... + 6 * 1/12 = \sum_{k=1}^{6} k * \mathbb{P}\{X = k\} = 2.5.$$

3. Random Walk

In this section we will discuss in detail the theorem of a Random Walk and the proof behind that theorem in the 1st and higher dimensions. Most of Section 3 follows from Lawler's Random Walk and the Discrete Heat Equation [2], with some foundational work set out by Spivak [3] and Levin [4], and details worked out by the author.

Theorem 3.1 (Random Walk in $\mathbb{Z}$). *: A random walker, with equal probability of moving ± one space, will pass through the origin infinitely many times with probability one.*

The distance, which is determined by the infinite amount of times our walker flips the coin, is composed of *independent random variables*, which are not reliant on the past according to Levin [4]. This process creates our random walk.

Our theorem says that after an infinite amount of coin tosses, our random walker will always end up back where he started, or the average value after an infinite amount of tosses will surely be zero, as long as the walker only moves in the 1st dimension. We model this random walk mathematically by defining, as mentioned in Definition 2.15, $S_n = \sum_{i=1}^{n} X_i$, where X_i are *independent random variables*, with $\{\mathbb{P} = 1\} = 1/2$, and $\{\mathbb{P} = -1\} = 1/2$.

We will show that $\mathbb{P}\{S_n = 0 \text{ for infinitely many n}\} = 1$.

Proof. [A random walker will pass through the origin infinitely many times]

Denote:

$$J_n = \begin{cases} 1 & S_n = 0 \\ 0 & otherwise. \end{cases}$$

Define V to equal the number of visits to the origin:

$$V = \sum_{n=0}^{\infty} J_{2n}.$$

The reason the number of visits to the origin takes the form of J_{2n}, is due to the fact that our walker can only return to the origin in an even number of steps. This is due to the fact that an odd number of steps will give us an odd integer that will never equal zero: In the case of one flip, we either have the values -1 or 1, in two flips we have the option of -2, 0, or 2 and in three flips, we have the options of -3,

-1, 1, or 3. This shows that an odd number of integers will give us an odd integer, and zero is not an odd integer.

Intuitively this is due to the fact that we can never go back to where we started unless we've undone the steps we've taken, which naturally doubles the amount of steps taken. Therefore in order to reach the origin, it necessitates that the total number of steps being expressed are a multiple of $2n$, which is in agreement with Spivak's definition of even numbers [3].

Since we want to confirm our average value is always zero, we need take the expected value of the visits to the origin:

$$E[V] = E[\sum_{n=0}^{\infty} J_{2n}].$$

Since J_{2n} will never be less than zero, if J_n converges it'll converge absolutely, therefore [3]:

$$E[\sum_{n=0}^{\infty}] = \sum_{n=0}^{\infty}[E].$$

Which means:

$$E[V] = \sum_{n=0}^{\infty} E[J_{2n}].$$

We want to take the expectation of J_{2n} when it is equal to 1, as we know that is the same as the expected number of visits when S_{2n} is equal to 0, or in other words how many times the random walker is expected to return to the origin. So we plug $X = J_{2n}$ and $z = 1$ into the definition of expectation (2.16), and we get:

$$E[J_{2n}] = \sum_{1} \mathbb{P}\{J_{2n} = 1\} = \mathbb{P}\{S_{2n} = 0\}.$$

Therefore:

$$E[V] = \sum_{n=0}^{\infty} E[J_{2n}] = \sum_{n=0}^{\infty} \mathbb{P}\{S_{2n} = 0\}.$$

Now we want to look at $\lim_{n\to\infty} \mathbb{P}\{S_{2n} = 0\}$, to figure out the behavior of the function near infinity.

First we need to determine what $\mathbb{P}\{S_{2n} = 0\}$ equals to. A key component to figuring out the probability of something is determining the total number of outcomes. When we think of a fair die, there are 6 faces, which are equally as likely to occur, so we divide 1 by 6 and we get the 1/6 to be the probability of getting any face.

In the case of random walk there are two options, Heads or Tails, -1 or +1. We've discussed earlier that $2n$ steps are needed to reach the origin, as it is necessary for there to be an even number of steps in order to get 0. Therefore there are 2^{2n} total possible combinations of +1 and -1 to add up to 0. But it isn't sufficient to just have $\mathbb{P}\{S_{2n} = 0\} = 1/2^{2n}$.

The statement above assumes that each combination of ± 1 would be unique. That when we have the event of flipping two heads and tails ($f = \{H, H, T, T\}$), that there would exist two unique events $f_1 = \{H_1, H_2, T_1, T_2\}$ and $f_2 = \{H_2, H_1, T_2, T_1\}$. But this notion is absurd, as there is no actual way to differentiate a head from a head, or a tail from a tail.

Therefore we need a way to insure this double counting doesn't occur, which is why we introduce the binomial coefficient $\binom{2n}{n}$ to our probability. After all, there are $2n!$ unique combinations of arranging heads and tails, such that the aforementioned absurdity occurs. So to prevent this occurrence we divide $2n!$ by the $n!$ ways to arrange all the heads and the $n!$ ways to arrange all the tails.

Since [3]:

$$\binom{N}{k} = \frac{N!}{(N-k)!k!},$$

it makes sense that we have $\binom{2n}{n}$ or $\frac{2n!}{n!n!}$ prevent our double counting from occurring.

Example 3.2. If you have the option of 2 heads and 2 tails you can arrange it in 6 unique ways: HHTT, TTHH, THTH,HTHT,HTTH,THHT. In this case n would be 2, and we have $4!/2!2! = 4 * 3 * 2/4 = 6$.

Therefore:

$$\mathbb{P}\{S_{2n} = 0\} = \binom{2n}{n} * \frac{1}{2^{2n}}.$$

Now we want to take:

$$\lim_{n\to\infty} \binom{2n}{n} * \frac{1}{2^{2n}} = \lim_{n\to\infty} \frac{2n!}{n!n!} * \frac{1}{2^{2n}}.$$

This limit is not obvious, so this is where Stirling's formula will be needed [2]. Stirling's formula shows that:

$$\lim_{n\to\infty} \frac{n!}{n^{(n+\frac{1}{2})}e^{-n}} = C.$$

Since the ratio equals some constant C, we can put C in the denominator of the original fraction and then we can replace $n!$ with $(Cn^{n+\frac{1}{2}}e^{-n})$. We will not solve for C in this paper, but Lawler goes through the process to determine that C is equal to $\sqrt{2\pi}$ [2]. Which means $n!$ is approximated by $\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}$.

Stirling's Formula. Define:

$$b_n = \frac{n!}{n^{n+\frac{1}{2}}e^{-n}}.$$

Note that $b_n > 0$ and $b_{n-1} > 0$, therefore $\frac{b_n}{b_{n-1}} > 0$.

We first want to show that:

$$\lim_{n\to\infty} b_n = C.$$

We can define:

$$b_n = b_1 * \frac{b_2}{b_1} * \frac{b_3}{b_2} * ... * \frac{b_{n-1}}{b_{n-2}} * \frac{b_n}{b_{n-1}} = b_1 * \prod_{j=2}^{n} \frac{b_j}{b_{j-1}}.$$

Since we are taking $\lim_{n\to\infty} b_n$, it'll be easier if we take $\lim_{n\to\infty} \log(b_n)$.

Note that [3]:

$$\log(b_n) = \log(b_1 * \prod_{j=2}^{n} \frac{b_j}{b_{j-1}}) = \log(b_1)+\log(\frac{b_2}{b_1})+\cdots+\log(\frac{b_n}{b_{n-1}}) = \log(b_1)+\sum_{j=2}^{n}\log(\frac{b_j}{b_{j-1}}).$$

Therefore:

$$\lim_{n\to\infty}\log(b_n) = \log(b_1) + \lim_{n\to\infty}\sum_{j=2}^{n}\log(\frac{b_j}{b_{j-1}}) = \log b_1 + \sum_{j=2}^{\infty}\log(\frac{b_j}{b_{j-1}}).$$

Note: Since we're splitting the limit across a finite sum: $\lim\sum = \sum\lim$, making the previous step valid [3].

Now if $\sum_{j=2}^{\infty}\log(\frac{b_j}{b_{j-1}})$ converges, then $\lim_{n\to\infty}\log(b_n)$ converges to some C, which is what we want. We can ignore b_1 since it is just e, some constant close to 3, so it won't impact our overall convergence.

In order to see if $\sum_{j=2}^{\infty}\log(\frac{b_j}{b_{j-1}})$ converges, it'll be easier if we define:

$$\delta_j = (\frac{b_j}{b_{j-1}}) - 1.$$

Therefore:

$$\sum_{j=2}^{\infty}\log(\frac{b_j}{b_{j-1}}) = \sum_{j=2}^{\infty}\log(1+\delta_j).$$

So if $\lim_{j\to\infty}\delta_j$ goes to zero fast enough, then $\sum_{j=2}^{\infty}\log(1+\delta_j)$ will converge.

Since the Taylor expansion of $\log(1+\delta_j)$ takes on the form similar to the alternating harmonic series, the starting value δ_j will always be the largest value of the expansion, multiplied by some constant (as each term added will never be larger than the term subtracted away, making the expansion always decrease in value).

Therefore $|\log(1+\delta_j)| \leq c{*}\delta_j$, so if $\sum_{j=2}^{\infty}\delta_j$ converges, then $\sum_{j=2}^{\infty}c{*}\delta_j$ also converges (as we just multiply that which it converges to by the constant c), and by the statement above $\sum_{j=2}^{\infty}\log(1+\delta_j)$ is finite [2].

Since we defined $\delta_j = \frac{b_j}{b_{j-1}} - 1$, in the case where $j = 1$, we get $b_1 = e$, but our expression for δ_j is only defined for $j \geq 2$.

It'll be easier to see if δ_j converges fast enough by looking at $\frac{b_j}{b_{j-1}}$. In that case:

$$\frac{b_j}{b_{j-1}} = \frac{\dfrac{j!}{j^{(j+1/2)} * e^{-j}}}{\dfrac{(j-1)!}{(j-1)^{(j-1+1/2)} * e^{(-(j-1))}}} = \frac{j! * (j-1)^{(j-1+1/2)} * e^{-(j-1)}}{(j-1)! * j^{(j+1/2)} * e^{-j}}$$

$$= \frac{j * (j-1)! * (j-1)^{j-1/2} * e^{-j} * e}{(j-1)! * j^{j+1/2} * e^{-j}} = \frac{j^{-(-1)} * (j-1)^{j-1/2} * e}{j^{j+1/2}}$$

$$= e * (\frac{j-1}{j})^{j-1/2} = e * (1 - \frac{1}{j})^{j}(1 - \frac{1}{j})^{(-1/2)}.$$

Recall that by the Taylor expansion rules [3]:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3.$$

The statement above means:

$$\log(1-x) = (-1)\int \frac{1}{1-x} = -x - \frac{1}{2} * x^2 - \frac{1}{3} * x^3.$$

As long as $|x| < 1$ then this Taylor expansion will converge, as it takes on the form of the alternating harmonic. This is due to the fact that the amount summed never exceeds the amount subtracted. Now we can arbitrarily choose $|x| < 1/2$, as $1/2$ is larger than zero but smaller than one, and a nice number overall [2].

There is a remainder left over when we subtract the Taylor expansion from the regular equation [3]:

$$F(x) - P_{n,a}(x) = R_{n,a}(x) = \frac{f^{n+1}(d)}{n+1!}(x-a)^{n+1}$$

Specifically we want, for some arbitrary constant K [2]:

$$|\log(1-x) - P_n(x)| \leq K * x^{n+1}.$$

This will be a favorable result as long as $|x| \leq 1/2$.

In our case we'll be looking at $x = \frac{1}{j}$, therefore the remainder term will be $K*(\frac{1}{j})^{k+1}$, where $k+1 \geq 2$. This means that the error will be small, so $R_k(\frac{1}{j}) < j^{-2}$. Since the error is so small, no matter how far the equation deviates from the Taylor Polynomial, by the comparison test we can still say that the whole sum converges. This means we can ignore the remainder term.

By plugging in $-x$ into the Taylor expansion of log(1-x), you get:

$$\log(1+x) = x - \frac{1}{2} * (x^2) + \frac{1}{3} * (x^3) + ...$$

So then plugging in $\frac{-1}{j}$ into the equation above you get:

$$\log(1 + \frac{-1}{j}) = \frac{-1}{j} - \frac{1}{2j^2} - \frac{1}{3j^3} + ...$$

Since we are looking at j large enough of $\frac{b_j}{b_{j-1}} - 1$, we can get a more favorable answer by taking the log of $\frac{b_j}{b_{j-1}}$:

$$\log(e*(1-\frac{1}{j})^{j}(1-\frac{1}{j})^{(-1/2)}) = \log(e)+\log(1-\frac{1}{j})^{j}+\log(1-\frac{1}{j})^{1/2} = \log(e)+j*\log(1-\frac{1}{j})+\frac{1}{2}*\log(1-\frac{1}{j}).$$

With the Taylor expansion of log in mind, now we look at j large enough of $|\log(\frac{b_j}{b_{j-1}})|$:

$$\leq 1 + j(-\frac{1}{j} - \frac{1}{2j^2} + \frac{C'}{j^3}) + -\frac{1}{2}(-\frac{1}{j} - \frac{1}{2j^2}) =$$

$$1 + (-1) + (-\frac{1}{2j}) + C'j^-2 + (\frac{1}{2j}) + \frac{1}{4j^2} = \frac{1}{4j^2} + C'j^-2 = \frac{C'}{j^2}.$$

Note that in a reasoning similar to ignoring the remainder term, we expand the polynomial that approximates $\log(1 + \frac{1}{j})$ only such that the final power will have j raised to the negative 2nd degree. Expanding further would be unnecessary, as j raised to a degree ≤ -2 will converge by the p-test in the final summation. Therefore we just need to look at the larger powers of j to make sure they don't add up to be larger than j^{-2} in the final summation, or our final summation won't converge.

Since $|\log(\frac{b_j}{b_{j-1}})| \leq \frac{C'}{j^2}$:

$$e^{\log(\frac{b_j}{b_{j-1}})} \leq e^{(\frac{C'}{j^2})}.$$

Therefore:

$$|\delta_j| = |\frac{b_j}{b_{j-1}} - 1| \leq e^{\frac{C'}{j^2}} - 1.$$

Then according to the Taylor expansion of e:

$$e^{\frac{C'}{j^2}} = 1 + \frac{C'}{j^2},$$

So:

$$|\delta_j| \leq \frac{C'}{j^2} - 1 + 1 = \frac{C'}{j^2}.$$

So there is some C' such that $|\delta_j| \leq \frac{C'}{j^2}$. Therefore:

$$\sum_{j=2}^{\infty} |\log(1 + \delta_j)| \leq \sum_{j=2}^{\infty} \delta_j \leq \sum_{j=2}^{\infty} |\delta_j| \leq \sum_{j=2}^{\infty} \frac{C'}{j^2}.$$

Recall the p-test, which says that [3]:

$$\sum_{n=0}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$.

Therefore the ratio converges to some unknown value C, which is different from the C' mentioned above, according to the p-test and the comparison test. This means we can replace $n!$ with $C * n^{n+1/2}e^{-n}$. □

Therefore:

$$\binom{2n}{n} * \frac{1}{2^{2n}} = \frac{2n!}{n!n!} * \frac{1}{2^{2n}} = \frac{(C * (2n)^{2n+1/2}e^{-2n})}{(C * n^{n+1/2}e^{-n})^2} * 2^{-(2n)} =$$

$$(\frac{C * (2n)^{2n+1/2}e^{-2n}}{C^2 * n^{2n+1}e^{-2n}}) * 2^{-(2n)} = (C" * 2^{2n+1/2}) * (\frac{n^{2n+1/2}}{n^{2n+1}}) * 2^{-(2n)}$$

$$= \frac{C\text{''} * \sqrt{2}}{\sqrt{n}} = \frac{C_0}{\sqrt{n}}.$$

For some constant C_0 [2]:

$$\lim_{n\to\infty} \frac{\mathbb{P}\{S_{2n} = 0\}}{(\frac{C_0}{\sqrt{n}})} = 1.$$

Which means whatever $\lim_{n\to\infty} \frac{C_0}{\sqrt{n}}$ approaches, that is what $\lim_{n\to\infty} \mathbb{P}\{S_{2n} = 0\}$ approaches.

Since the values they have are the same near infinity, we can replace $\mathbb{P}\{S_{2n} = 0\}$ with $\frac{C_0}{\sqrt{n}}$ and see if the answer has the ideal result of diverging.

Therefore, according to the p-test:

$$\sum_{n=0}^{\infty} \frac{C_0}{\sqrt{n}} = \infty.$$

In this case $p = 1/2$, which is less than 1, which causes the sum to diverge by the p-test. Since we can pull out the constant, C_0 has no affect on the final summation.

Therefore $E[V] = \infty$, which means the random walker is expected to return to the origin infinitely many times. □

Proof. [The probability will be one]

Denote q to be the probability that the random walker passes through the origin infinitely many times. We want to show that $q = 1$, so let us assume not. Suppose $q < 1$ instead.

If $q < 1$, the probability that $\mathbb{P}\{V = 1\} = 1 - q$, because $V = 1 \iff q = 0$, or the random walker only starts at the origin and never returns afterwards.

Let us look at some examples to determine the general case. Suppose we have a fair coin, with probability 1/2 that we will get heads or tails. Let us suppose we have one flip to get heads, the probability that this will occur is 1/2. Suppose we now want to get two heads, and we only have two flips to achieve this. We get 1/2 probability from the first flip, multiplied by 1/2 probability from the second flip. For the third flip we multiply by another 1/2 and so on. In this case $(1 - q)$ is 1/2, and then we multiply by q each time the amount of heads we want increases.

Now in the case with a six headed die, say we want to roll and get the number 3. If we get a 3, our random walker will never return to the origin. But if we roll any other number, our random walker will return to the origin. So suppose $V = 1$, that means after our first roll we will never return to the origin. This result relies on the probability we will get a 3 in one roll, which is 1/6. Now suppose we want $V = 2$, which means we've visited the origin in two rolls: We've started at the origin, with one roll we've got something that isn't a 3, returning us to the origin, and the second roll we get a 3. This relies on the probability we will get something that isn't 3 on our first roll, which is 5/6 and then the next roll you get the probability of getting 3 which is 1/6. Therefore $E[V = 2] = 1/6 * 5/6$ Now with the increase of number of visits, we multiply by the probability of not getting a 3, which is 5/6. Therefore $E[V = j] = 1/6 * (5/6)^{j-1}$.

In general [2]:

$$\mathbb{P}\{V = j\} = (1-q) * q^{j-1} \ for \ j = 1, 2, 3...$$

Plugging the prior result into our definition of expectation, we find:

$$E[V] = \sum_{j=1}^{\infty} j * \mathbb{P}\{V = j\} = \sum_{j=1}^{\infty} j * (1-q) * q^{j-1} = \sum_{j=1}^{\infty} j * q^{j-1} - j * q^{j}.$$

By expanding the result, we get:

$$1 - q + 2(q) - 2(q^2) + 3(q^2) - 3(q^3) + 4(q^4)...,$$

which simplifies to the regular geometric series:

$$1 + q + q^2 + q^3 + ...,$$

Therefore:

$$E[V] = \sum_{j=1}^{\infty} q^j = \lim_{n\to\infty} \sum_{j=1}^{n} q^j.$$

Let:

$$E[V] = \lim_{n\to\infty} S_n,$$

where:

$$S_n = \sum_{j=1}^{n} q^j.$$

Multiplying S_n by $(q-1)$, we get:

$$S_n(q-1) = \sum_{j=1}^{n} q^j(q-1) = 1 + q^{n+1}.$$

Then dividing the above by $(q-1)$, we get:

$$S_n = \frac{1 + q^{n+1}}{q-1}.$$

We've assumed $q < 1$, so when we take the limit:

$$\lim_{n\to\infty} S_n = \frac{1}{1-q} = E[V].$$

In this case $E[V] = \frac{1}{1-q} \neq \infty$, which contradicts part one. Therefore $q = 1$.

□

Theorem 3.3 (Random Walk in $\mathbb{Z}$ in dimensions 3 or higher)**.** *: A random walker, with equal probability moving ± one space throughout each dimension with dimension higher than 3, will pass through the origin finitely many times, or infinitely many times with probability zero.*

Proof. In the multi-dimension case according to Lawler [2]:

$$\lim_{n\to\infty} \frac{\mathbb{P}\{S_{2n}=0\}}{C_0/n^{d/2}} = 1.$$

Therefore according to the p-rule mentioned above, if $d > 2$ then $p > 1$, which means that $\mathbb{P}\{S_{2n} = 0\}$ converges. Therefore the random walker passes through the origin only finitely many times. □

These results imply that if the walker is in the 1st and 2nd dimension, the walker will always return to where they started, but as soon as higher dimensions are introduced, they will not. It may seems counter intuitive at first, as one could think since each step forward and back has equal probability, that despite the change of the number of dimensions, the walker would have as equal likely probability of ending up where they started. It would seem that in all cases the walker either returns to the origin finitely or infinitely many times. But with further investigation, the addition of more axis means there are too many combinations of random variables. There are too many options of back and forth with each step. This makes it impossible to always backtrack and end up back where one started, infinitely many times. One may be lucky and end up back where they started, but only finitely many times.

Since the flip of a coin only has two possibilities of equal probability, and after 2 dimensions it becomes impossible to visit the origin infinitely many times, it raises questions of what would occur with deviation. What would occur if the Random Walk was based on the probability of a die, or a die that was weighted, such that the probabilities wouldn't be the same? I presume the random walker would end up back to the origin only finitely many times in lower dimensions, or even not at all. Maybe if we changed the experiment such that the random walker would return to the die's respective average values, instead of the origin, it may occur an infinite amount of times in the lower dimensions.

Acknowledgments

I would like to thank my mentor Isabella for her support and the countless hours spent guiding me. Thank you to Professor Lawler for his time and guidance, to Professor Fefferman who inspired my curiosity about probability, and Peter May for his inclination. And a special thanks to my sister who was just as excited as I was about this learning opportunity.

4. Bibliography

References

[1] David Williams. Probability with Martingales. Cambridge University Press. 2008. Pages 15-19,23, 24, 29, 31, 38.
[2] Gregory Lawler. Random Walk and the Heat Equation. American Mathematical Soc. 2011. http://www.math.uchicago.edu/ lawler/reu.pdf. Pages 5-17.
[3] Michael Spivak. Calculus Fourth Edition. Publish or Perish. 2008. Pages 25, 178, 179, 421-423, 449, 472, 474,475, 479-81, 515.
[4] David Asher Levin. Markov Chains and Mixing Times. American Mathematical Soc. 2009. Pages 30-33.
[5] http://physics.bnu.edu.cn/application/faculty/tuzhanchun/biophys/L08.pdf