Combinatorics, Probability and Computing (2004) 00, 000–000. © 2004 Cambridge University Press
DOI: 10.1017/S0000000000000000 Printed in the United Kingdom

Simulating a Random Walk with Constant Error

JOSHUA N. COOPER† and JOEL SPENCER

Department of Mathematics, Courant Institute of Mathematics, New York University,
251 Mercer St, New York, NY 10012–1185
cooper@cims.nyu.edu and spencer@cs.nyu.edu

We analyze Jim Propp's P-machine, a simple deterministic process that simulates a random walk on $\mathbb{Z}^d$ to within a constant. The proof of the error bound relies on several estimates in the theory of simple random walks and some careful summing. We mention three intriguing conjectures concerning sign-changes and unimodality of functions in the linear span of $\{p(\cdot, \mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\}$, where $p(n, \mathbf{x})$ is the probability that a walk beginning from the origin arrives at $\mathbf{x}$ at time n.

Consider the following machine, which we call the P-machine after its progenitor, Jim Propp. We begin with a set of chips on "even" vertices of $\mathbb{Z}^d$, i.e., those an even L^1-distance from the origin. We label by $\mathbb{Z}_{2d} = \{0, 1, \ldots, 2d-1\}$ the $2d$ directions of Z^d. (For example: North, South, East and West for $d = 2$.) On each vertex $\mathbf{v}$ there will be a "rotor" which will have a state $j \in \mathbb{Z}_{2d}$. The initial states of the rotors can be set arbitrarily. The rotors work as follows: when one feeds it a chip, a rotor changes to state $j+1$ (addition in $\mathbb{Z}_{2d}$) and sends the chip to the vertex in direction j. Each time unit we feed all of the chips currently on each vertex to the rotor on that vertex. The result is a new distribution of chips. We imagine doing this operation for total time n.

Now, consider the following process. We again begin with a set of chips on "even" vertices of $\mathbb{Z}^d$. Each chip takes an n-step simple random walk from its starting point. At time n, we expect that the P-machine and the random process should give rise to similar distributions if they begin with the same configuration of chips. For the random walk process we may consider the expected number of chips that will be at $\mathbf{v}$ at time n. Our main result is that the difference between this expected number and the actual number at $\mathbf{v}$ at time n in the deterministic P-machine is bounded uniformly – *irrespective of how much time has passed, what the original chip distribution was, the starting states of the rotors, or even the choice of* $\mathbf{v}$*!*

As an example, suppose n is even, $d = 1$, we begin with n chips at position 0, and the total time is n. The random walk model will have an expected number $n\binom{n}{n/2}2^{-n} =$

† Research supported by NSF Grant DMS-0303272.

$\Theta(\sqrt{n})$ chips at position $\mathbf{v} = 0$. The deterministic P-machine will give that number with only constant error.

As a further interpretation, consider a "linear machine" in which a pile of c chips are split evenly, with $\frac{c}{2d}$ going in each direction. The final number (which, in general, will not be an integer) of chips at a position will simply be the expected number of chips in the random walk machine. We may consider the linear machine a "relaxation" of the P-machine, and our result states that this relaxation produces only constant "error."

1. The Proof

We introduce some notation. Chip distributions will be written as functions χ from $\mathbb{Z}^d$ to the nonnegative integers. We let n represent the total time of the process and let t represent the *time remaining* in the process so that t counts backwards from n to 0. Let χ_t denote the configuration of the P-machine with time remaining t. Thus χ_n is the initial configuration and χ_0 is the final configuration. We prescribe an ordering $\mathbf{e}_0, \ldots, \mathbf{e}_{2d-1}$ on those vectors (directions) with all coordinates zero except for one which is $+1$ or -1. We fix a map $\sigma : \mathbb{Z}^d \to \mathbb{Z}_{2d}$ denoting the initial settings of the rotors. Define S_t, a map from $\mathbb{Z}^d$ to $\mathbb{R}$, by setting $S_t(\mathbf{v})$ to be the probability that a random walk of time t beginning at the origin ends at $\mathbf{v}$. For any distribution χ and any t we define

$$W_t(\mathbf{v}, \chi) = \sum_{\mathbf{w} \in Z^d} \chi(\mathbf{w}) S_t(\mathbf{v} - \mathbf{w}), \tag{1.1}$$

This W_t acts as a weighted sum of all the chips when t is the time remaining. Thus $W_0(\mathbf{v}, \chi_0)$ is the actual number of chips at $\mathbf{v}$ at the conclusion of the P-machine process whereas $W_n(\mathbf{v}, \chi_n)$ is the expected number of chips at $\mathbf{v}$ at the conclusion of the random walk process.

We (following [1]) write $\mathbf{v} \leftrightarrow n$ if the sum of the coordinates of $\mathbf{v}$ has the same parity as n, and $\mathbf{w} \leftrightarrow \mathbf{w}$ if $\mathbf{v}$ and $\mathbf{w}$ have the same parity in this sense. Call a configuration χ "even" ("odd") if, for all $\mathbf{v}$ in the support of χ, $\mathbf{v} \leftrightarrow 0$ (resp., $\mathbf{v} \leftrightarrow 1$).

Our main result is the following.

Theorem 1.1. *There is a constant C_d, depending on the dimension d but not the initial even (or odd) configuration χ_n, the total time n, nor the initial rotor settings σ so that*

$$|W_0(\mathbf{v}, \chi_0) - W_n(\mathbf{v}, \chi_n)| < C_d.$$

Proof. Clearly, we may assume without loss of generality that the configuration is even. Then, for $n \geq t \geq 0$ we define

$$X_t = W_t(\mathbf{v}, \chi_t) \tag{1.2}$$

We may interpret X_t as the expected number of chips at $\mathbf{v}$ at the conclusion of a process in which the first $n - t$ steps are by the deterministic P-machine and the remaining t steps are in the random walk model. The sequence X_t, $n \geq t \geq 0$, is then reminiscent of a martingale. We first expand "in time":

$$X_n - X_0 = \sum_{t=1}^{n} X_t - X_{t-1} = \sum_{t=1}^{n} W_t(\cdot, \chi_t) - W_{t-1}(\cdot, \chi_{t-1}) \tag{1.3}$$

The summand in (1.3) is the change in weight during one step of the P-machine. To analyze this change we set

$$T^j(\mathbf{w}, t) = S_t(\mathbf{w}) - S_{t-1}(\mathbf{w} - \mathbf{e}_j) \tag{1.4}$$

This represents the change in weight of a single chip that was at position $\mathbf{v} + \mathbf{w}$ with time remaining t and with the rotor at that position set to state j. Observe, critically, $S_t(\mathbf{w})$ is the average of all possible $S_{t-1}(\mathbf{w} - \mathbf{e}_j)$ so that

$$\sum_{j=0}^{2d-1} T^j(\mathbf{w}, u) = 0 \tag{1.5}$$

Let $\Delta_{\mathbf{w}}$ denote the total contribution to $X_n - X_0$ from all chips moved from position $\mathbf{v} + \mathbf{w}$. (Effectively we are here moving the origin to $\mathbf{v}$.) Suppose that during the process a chip was moved from $\mathbf{v} + \mathbf{w}$ a total of M times and that the i-th time occured (starting the count at zero) when the remaining time was t_i. Let c denote the initial state of position $\mathbf{v} + \mathbf{w}$. Then

$$\Delta_{\mathbf{w}} = \sum_{i=0}^{M-1} T^{c+i}(\mathbf{w}, t_i) \tag{1.6}$$

(As before, we consider state $c + i$ as an element of $\mathbb{Z}_{2d}$.) Reversing the sum of (1.3) we express

$$X_n - X_0 = \sum_{\mathbf{w} \in \mathbb{Z}^d} \Delta_{\mathbf{w}}. \tag{1.7}$$

1.1. Dimension One

The case of $d = 1$ is somewhat special, in that it requires separate arguments and is also susceptible to a more precise analysis. Nonetheless, it closely resembles the higher dimensional cases, and so we produce the proof in the one-dimensional case first.

All vector quantities are simply integers when $d = 1$, and we may write

$$T^j(w, t) = S_t(w) - S_{t-1}(w \pm 1)$$

where the choice of sign is $(-1)^j$. Since, by (1.5), $T^0(w, t) = -T^1(w, t)$, Δ_w can be written as an alternating sum of $T^1(w, t_i)$ for some nonincreasing sequence t_i.

Now, we wish to show that the sequence $\{T^j(w, t)\}_{t \geq 1}$ is unimodal. The unimodality of the sequence $T^1(w, t)$ is important because it implies that $T^1(w, t)$ is the concatenation of two monotone subsequences. An alternating sum of any monotone sequence which has all the same sign is bounded by its largest term. Therefore, we may estimate Δ_w by twice the maximum of $T^1(w, t)$.

It clearly suffices to consider the cases of $w \geq 1$ and $w = 0$ – otherwise write Δ_w in terms of T^0 instead of T^1. Then, for $t \leftrightarrow w$ with $w \geq 1$,

$$
\begin{aligned}
T^1(w,t) &= 2^{-t}\binom{t}{(t+w)/2} - 2^{-t+1}\binom{t-1}{(t+w)/2-1} \\
&= 2^{-t+1}\binom{t-1}{(t+w)/2-1}\left(\frac{-w}{t+w}\right),
\end{aligned} \tag{1.8}
$$

which is always negative. To determine when $T^j(w,t)$ is increasing or decreasing, we consider the ratio of consecutive terms in the range $t \geq w$:

$$
\begin{aligned}
\frac{T^1(w,t+2)}{T^1(w,t)} &= \frac{t+w}{4(t+w+2)} \cdot \frac{4t(t+1)}{(t+w)(t-w+2)} \\
&= \frac{t(t+1)}{(t+2)^2-w^2}.
\end{aligned}
$$

When $t(t+1) \leq (t+2)^2 - w^2$, i.e., when $t \geq (w^2-4)/3$, the quantity $T^1(w,t)$ is increasing; when $t \leq (w^2-4)/3$ it is decreasing. (For $t < w$, the sequence is identically 0.) Hence, $T^1(w,t)$ is unimodal for those t with $t \leftrightarrow w$.

To compute the maximum of $T^1(w,t)$, we plug in $t = w^2/3 + O(1)$, getting $|T^1(w,t)| = O(w^{-2})$. So, any alternating sum of $T^1(w,t_i)$ for t_i nonincreasing satisfies this same bound. (It is easy to see that, in the case of $w = 0$, $T^1(w,t) = 2^{-t-1}\binom{t}{t/2}$ is also unimodal and bounded by a constant.) Therefore, $\sum_{w\in\mathbb{Z}} O(w^{-2})$ converges, and Theorem 1.1 follows in the case of $d = 1$, by (1.7).

1.2. Higher Dimensions

In dimensions $d \geq 2$, the situation is more complicated. For $d \geq 3$ such elegant formulas as (1.8) simply do not exist, and we must use estimates for the terms in question. We suppress an argument for $d = 2$ using exact formulae as our methods below apply to all $d \geq 2$.

Claim 1. $\Delta_{\mathbf{w}} = O(|\mathbf{w}|^{-(d+1)} \ln^{d-1}|\mathbf{w}|)$. *That is, there exists a constant K (depending only on the dimension) such that $|\Delta_{\mathbf{w}}| \leq K|\mathbf{w}|^{-(d+1)}\ln^{d-1}|\mathbf{w}|$ for any choice of sequence $t_0, \ldots, t_{M-1}$ and initial state c and any $\mathbf{w} \in \mathbb{Z}^d$ with sufficiently high norm. Further, for any fixed $\mathbf{w}$ there is a constant bound on $|\Delta_{\mathbf{w}}|$, independent of the choice of sequence $t_0, \ldots, t_{M-1}$.*

The case when $\mathbf{w}$ is fixed is relatively simple and is given by Claim 6.

Consider the expansion (1.6) of $\Delta_{\mathbf{w}}$. For each i with $c + i = 0$, apply (1.5) and replace $T^0(\mathbf{w}, t_i)$ with $\sum_{j\neq 0} -T^j(\mathbf{w}, t_i)$. Now for each state $j \neq 0$ consider the sum of all terms with T^j. When $c + i = j$ these will be $T^j(\mathbf{w}, t_i)$ and when $c + i = 0$ these will be $-T^j(\mathbf{w}, t_i)$. These signs will alternate, so that the sum can be written

$$T^j(\mathbf{w}, t_1) - T^j(\mathbf{w}, t_2) + \ldots \pm T^j(\mathbf{w}, t_s) \tag{1.9}$$

where the $t_1, t_2, \ldots$ form a nonincreasing sequence. (The initial sign might be negative, but that will not matter.) Removing copies of the same term we can assume the $t_1, t_2, \ldots$

form a decreasing sequence. (In particular, suppose the number of chips on a given position at a given time is divisible by $2d$. Then for each state j the positive and negative contributions will cancel out. This is to be expected, as when the chips are distributed evenly the weight function will not change.) We have expressed $\Delta_{\mathbf{w}}$ as the sum of a constant $(2d-1)$ of these sequences so that to show Claim 1 it suffices to show

Claim 2. $T^j(\mathbf{w}, t_1) - T^j(\mathbf{w}, t_2) + \ldots \pm T^j(\mathbf{w}, t_s) = O(|\mathbf{w}|^{-(d+1)} \ln^{d-1} |\mathbf{w}|)$ *and is bounded by a constant for any fixed* $\mathbf{w}$.

As a further reduction we expand

$$T^j(\mathbf{w}, t) = S_t(\mathbf{w}) - S_{t-1}(\mathbf{w} - \mathbf{e}_j) = \frac{1}{2d} \sum_k S_{t-1}(\mathbf{w} - \mathbf{e}_k) - S_{t-1}(\mathbf{w} - \mathbf{e}_j)$$

where k ranges over the $2d$ states. The sum of Claim 2 splits into $2d-1$ sums. (When $k = j$ we get zero.) For fixed distinct states j, k we set $\mathbf{x} = \mathbf{w} - \mathbf{e}_j$ and $\mathbf{z} = \mathbf{e}_j - \mathbf{e}_k$ and, again following [1],

$$\nabla_{\mathbf{z}}(\mathbf{x}, t) = S_t(\mathbf{x} + \mathbf{z}) - S_t(\mathbf{x})$$

As $|\mathbf{x}|, |\mathbf{w}|$ differ by at most one they are asymptotically the same and so Claim 2 reduces to showing

Claim 3. $\nabla_{\mathbf{z}}(\mathbf{x}, t_1) - \nabla_{\mathbf{z}}(\mathbf{x}, t_2) + \ldots \pm \nabla_{\mathbf{z}}(\mathbf{x}, t_s) = O(|\mathbf{x}|^{-(d+1)} \ln^{d-1} |\mathbf{x}|)$ *and is bounded by a constant for any* $\mathbf{x}$.

This we shall show for any fixed $\mathbf{z} \leftrightarrow 0$ and any decreasing sequence $t_1, \ldots, t_s$.

First, we show that the contribution from small t (i.e., near the conclusion of the process) is small. With foresight we set

$$T = K|\mathbf{x}|^2 / \ln^2(|\mathbf{x}|) \tag{1.10}$$

where K is a large constant to be specified. T will be our cutoff between small and not small. A result from [1] will be useful here: For some constant $c > 0$, a random walk beginning at the origin ends at a point at least a distance $\alpha t^{1/2}$ away at time t with probability at most $ce^{-\alpha}$. Now, $|\nabla_{\mathbf{z}}(\mathbf{x}, t)|$ is certainly bounded by the probability that the walk ends at least a distance $|\mathbf{x}|$ from the origin plus the probability that it ends at least $|\mathbf{x} + \mathbf{z}|$ from the origin, so

$$\begin{aligned} |\nabla_{\mathbf{z}}(\mathbf{x}, t)| &\leq ce^{-|\mathbf{x}|/t^{1/2}} + ce^{-|\mathbf{x}+\mathbf{z}|/t^{1/2}} \\ &\leq 2ce^{-(|\mathbf{x}|-|\mathbf{z}|)/t^{1/2}}. \end{aligned}$$

Now consider the the contribution from the small $t < T$. We select K so that all $|\nabla_{\mathbf{z}}(\mathbf{x}, t)| = O(|\mathbf{x}|^{-d-3})$. The total contribution to the sum of Claim 3 from all $t \leq T$, there being fewer than $|\mathbf{x}|^2$ terms, is $O(|\mathbf{x}|^{-d-1})$.

We can now assume that all $t_i \geq T$. We approximate the $\nabla_{\mathbf{z}}(\mathbf{x}, t)$ by the following special case of the "Local Central Limit Theorem" for simple random walks (q.v. [1]):

$$|\nabla_{\mathbf{z}}(\mathbf{x},t) - (p(t,\mathbf{x}+\mathbf{z}) - p(t,\mathbf{x}))| = |\mathbf{x}|^{-2} O(t^{-(d+1)/2}) \tag{1.11}$$

where $|\cdot|$ denotes the L^2-norm, $t \leftrightarrow \mathbf{x}$, and

$$p(t,\mathbf{x}) := 2(\frac{d}{2\pi t})^{d/2} e^{-d|\mathbf{x}|^2/2t} \tag{1.12}$$

(Here, and throughout the rest of the proof, all hidden constants may depend on $\mathbf{z}$. This does not affect our original problem since $\mathbf{z}$ takes on only a finite set of values.)

An alternating sum of $\nabla_{\mathbf{z}}(\mathbf{x},t)$ is then, in absolute value, at most the absolute value of the alternating sum of the approximation $p(t,\mathbf{x}+\mathbf{z}) - p(t,\mathbf{x})$ and the sum of the absolute values of the errors. The sum of the absolute values of the errors for $t \geq T$ is then $O(|\mathbf{x}|^{-2} T^{(-d+1)/2})$. (Note that this argument would *not* work for $d = 1$ as the bounds on the errors would form a divergent harmonic series.) Our choice of cutoff T ensures that this is $O(|\mathbf{x}|^{-(d+1)} \ln^{d-1} |\mathbf{x}|)$.

Now, we wish to show that an alternating sum of the quantity $f(t) = p(t,\mathbf{x}) - p(t,\mathbf{x}+\mathbf{z})$ is $O(|\mathbf{x}|^{-(d+1)})$. We regard this as the "main term" of our calculations. This is the critical place where we need the alternation of the sum. That the sum alternates is a direct consequence of the construction of the P-machine, and represents in our minds the "cancelling out process" when a single rotor distributes chips in a (relatively) even manner.

Claim 4. $f(t) = O(|\mathbf{x}|^{-(d+1)})$

Let $\rho = |\mathbf{x}+\mathbf{z}| - |\mathbf{w}|$, so $|\rho| \leq |\mathbf{z}|$. Since $f(t)$ is zero when $\min\{|\mathbf{x}|, |\mathbf{x}+\mathbf{z}|\} > t$, we may write

$$\begin{aligned} f(t) &\ll t^{-d/2} |e^{-d|\mathbf{x}|^2/2t} - e^{-d|\mathbf{x}+\mathbf{z}|^2/2t}| \\ &= t^{-d/2} e^{-d|\mathbf{x}|^2/2t} |1 - e^{-d\rho(|\mathbf{x}+\mathbf{z}|+|\mathbf{x}|)/2t}| \\ &\ll t^{-d/2} e^{-d|\mathbf{x}|^2/2t} d|\rho|(|\mathbf{x}+\mathbf{z}| + |\mathbf{x}|)/2t \\ &\ll t^{-d/2-1} |\mathbf{x}| e^{-d|\mathbf{x}|^2/2t}. \end{aligned}$$

It is a matter of elementary calculus to show that this function is maximized when $t = \frac{d}{d+2}|\mathbf{x}|^2$, in which case $f(t)$ satisfies Claim 4.

Claim 5. *The sequence $f(t)$ has at most 6 local extrema. (We say that t is a local extremum if $f(t) > \max[f(t+2), f(t-2)]$ or $f(t) < \min[f(t+2), f(t-2)]$).*

Letting $B = 2(d/2\pi)^{d/2}$,, we may write

$$p(t,\mathbf{x}) - p(t,\mathbf{x}+\mathbf{z}) = Bt^{-d/2}(e^{-d|\mathbf{x}|^2/2t} - e^{-d(|\mathbf{x}+\mathbf{z}|^2/2t})$$

Setting the derivative equal to zero gives

$$e^{-\gamma y} = K_1 + \frac{K_2}{y-1} \tag{1.13}$$

after making the substitutions $R = |\mathbf{x}|^2$, $S = |\mathbf{x}+\mathbf{z}|^2$, $y = R/t$, $K_1 = S/R$, $K_2 = (S-R)/R$, and $\gamma = (R-S)/2dR$. This splits into the two regions $y < 1$ and $y > 1$. In each region the number of zeroes is at most one more than the number of zeroes of the derivative. Therefore, setting the derivative of (1.13) equal to zero gives the equation

$$e^{-\gamma y}(y-1)^2 - K_3 = 0$$

The number of zeroes of this is at most (in each region) the number of zeroes of its derivative, which is

$$e^{-\gamma y}[2(y-1) - \gamma(y-1)^2] = 0.$$

This equation clearly has at most two solutions, and the proof of Claim 5 is complete.

From Claim 5 the sequence $f(t)$ can be split into a bounded (≤ 7) number of monotone sequences. *On each such sequence an alternating sum is at most twice the maximal absolute value of the terms.* Applying Claim 4 the entire sum is $O(|\mathbf{x}|^{-(d+1)})$ which concludes Claim 3 and thus Claim 1 in the asymptotic case when $\mathbf{x} \to \infty$.

Claim 6. *For any fixed* $\mathbf{w}$ *there is a constant bound on* $|\Delta_{\mathbf{w}}|$.

The reductions to Claim 3 are as given. The sequence $p(t, \mathbf{x}+\mathbf{z}) - p(t, \mathbf{x})$ clearly has absolute value at most two and the argument of Claim 5 holds for any $\mathbf{x}$ so an alternating sum of the $p(t, \mathbf{x}+\mathbf{z}) - p(t, \mathbf{x})$ is bounded. The error terms, from (1.11), are $O(t^{-(d+1)/2})$ and so their sum is bounded for all $d \geq 2$.

Having shown Claim 1, we may now show Theorem 1.1. Equation (1.7) and this Claim gives

$$X_n - X_0 = O\Big(\sum_{\mathbf{w} \in Z^d} |\mathbf{w}|^{-(d+1)} \ln^{d-1} |\mathbf{w}|\Big) = O(1). \tag{1.14}$$

2. Conjectures

In the course of proving the main theorem, we encountered the following intriguing conjecture.

Conjecture 1. *For each* $\xi : \mathbb{Z}^d \to \mathbb{R}$ *with finite support, and for all* $\mathbf{v} \in \mathbb{Z}^d$*, the function* $W_{2n}(\mathbf{v}, \xi)$ *has at most* K *sign changes in the variable* n*, where* K *depends on* ξ *and* d *but not on* $\mathbf{v}$*.*

To see where this statement comes from, note that, for any ξ, we may always use the harmonic property to write $W_{2n+2}(\mathbf{v}, \xi) - W_{2n}(\mathbf{v}, \xi) = W_{2n}(\mathbf{v}, \xi')$ for some $\xi' : \mathbb{Z}^d \to \mathbb{R}$. Then, $W_{2n}(\mathbf{v}, \xi)$ has a bounded number of local extrema if and only if $W_{2n}(\mathbf{v}, \xi')$ has a bounded number of sign changes. Therefore, if this conjecture were true, we could simply estimate the alternating sums in the proof by a constant times their largest term, and the burden of bounding the error terms in (1.11) would be reduced somewhat. Our intuition is that this statement is true because "oscillations" in $p(n, \mathbf{v})$ occur only to the modulus 2.

Furthermore, we believe that the polylogarithmic factor in the bound on $|\Delta_{\mathbf{w}}|$ is unnecessary. Indeed, this would follow from Conjecture 1.

Conjecture 2. *For any d, $|\Delta_{\mathbf{w}}| = O(|\mathbf{w}|^{-(d+1)})$.*

For very simple ξ, such as one whose support consists of a single vertex, we actually expect unimodality. That is,

Conjecture 3. *The probability that a random walk beginning at the origin in Z^d arrives at $\mathbf{v}$ at time $n \leftrightarrow \mathbf{v}$ is a unimodal function of n.*

The estimates given above are inadequate to prove this simple statement. Furthermore, attempting to show unimodality via log-concavity is futile, since these functions are often (perhaps always?) log-convex in the tail. In the $d = 1$ case, we can show unimodality by brute force. Indeed, $p(n, x) = 2^{-n}\binom{n}{(n+x)/2}$ for $x \geq 0$, and the ratio of consecutive terms is given by

$$\frac{p(n, x)}{p(n-2, x)} = \frac{n(n-1)}{n^2 - x^2}$$

so the sequence is 0 for $n < x$, increasing up to $n = x^2$, and decreasing thereafter. A similar argument also works for $d = 2$. It is clear that it suffices to prove unimodality for the half-quadrant where $x \geq 0$ and $y \leq x$. Note that, in a random walk on $\mathbb{Z}^2$, the quantities $x + y$ and $x - y$ vary *independently* at each step. Furthermore, $x + y$ and $x - y$ follow one-dimensional simple random walks. Therefore, we may decompose the two-dimensional walk as follows:

$$\begin{aligned} p(n, (x, y)) &= p(n, x + y) \cdot p(n, x - y) \\ &= 4^{-n}\binom{n}{(n + x + y)/2}\binom{n}{(n + x - y)/2}. \end{aligned}$$

If we consider the ratio of successive terms,

$$\frac{p(n, (x, y))}{p(n-2, (x, y))} = \frac{n^2(n-1)^2}{(n + x + y)(n + x - y)(n - x + y)(n - x - y)}.$$

In the range $n > x + y$, this quantity is ≤ 1 precisely when $g(n) = 2n^3 - rn^2 + s$ is positive, where $r = 2x^2 + 2y^2 + 1$ and $s = (x^2 - y^2)^2$. When $n = x + y$,

$$g(x + y) = -(x + y)^2(x + y - 1)^2 \leq 0.$$

On the other hand, g' has two roots: one at $n = 0$ and one at $n = r/3$. Since $0 \leq x + y$, at most one root of g occurs at in $[x + y, +\infty)$, so we may conclude that $p(n, (x, y))$ is unimodal in n.

Unfortunately, this kind of analysis does not seem to be possible in higher dimensions, so further insights are necessary to resolve the conjecture when $d \geq 3$.

3. Acknowledgements

Thank you to Noam Berger, Jim Propp, and David Wilson for helpful discussions and insights.

References

[1] G. F. Lawler, Intersections of random walks, *Probability and its Applications*, Birkhäuser, Boston, MA, 1991.