ON A SELF-AVOIDING RANDOM WALK

by

Thomas Kuczek and Keith Crank[1]
Department of Statistics, Purdue University, West Lafayette, IN, USA
Owen Graduate School of Management, Vanderbilt University, Nashville, TN, USA

Technical Report # 87-31

Department of Statistics
Purdue University

Revised June 1990

[1]This work was supported by The Dean's Fund for Research of the Owen Graduate School of Management of Vanderbilt University.

△

ON A SELF-AVOIDING RANDOM WALK

by

Thomas Kuczek
Department of Statistics
Purdue University
West Lafayette, IN, USA

and

Keith Crank
Owen Graduate School of Management
Vanderbilt University
Nashville, TN, USA

ABSTRACT

A self-avoiding random walk on Z^1 is considered. By conditioning on certain points with regeneration type properties called "break points", it is shown that the set of occupied points grows in a linear fashion. The utility of break points is that they greatly simplify the conditioning involved in studying the process.

Key words: random walk, break points

AMS 1980 subject classification: Primary 60K35

1. Introduction

The behavior of a stochastic process is often difficult to study owing to the complexities of the process as it evolves over time. In this paper we are interested in studying the behavior of such a process, a self-avoiding random walk. As defined in Spitzer (1954), a random walk is self-avoiding if we condition on those paths which are self-avoiding, i.e. which do not step to a point previously occupied. In order to study the process, we show the existence of points with regeneration-type properties termed "break points". Break points occur at times which are not stopping times, but which have stopping time like properties, which allow one to decompose edge behavior into independent pieces. One can then use renewal type arguments such as in Murphree (1987) who studied a related problem of the transient renewal process. Break points have more general applicability. In an entirely different setting, they were used to study the edge behavior of oriented percolation (Kuczek (1989)).

Section 2 contains definitions and notation. In Section 3 break points for this process are defined and shown to give the independent structure to the process. In Section 4 the limit behavior of the process is studied. Section 5 contains some further results relevant to the study of the process, while Section 6 has some concluding comments.

2. Definitions and Notation

Let $X_1, X_2, \ldots$, be a sequence of independent, identically distributed random variables, which are integer valued with distribution function $G(\cdot)$. Also, assume that the distribution of the random variable X_1 is symmetric about 0, has no mass at $\{0\}$, and does not have all its mass on $\{-1, 1\}$. Let

$$S_0 = 0$$

and

$$S_n = \sum_{i=1}^{n} X_i$$

denote the partial sums of the sequence $\{X_i\}$. To describe and study the self-avoiding paths, we introduce the following notation.

$$\Omega_n = \{(0, x_1, \ldots, x_n) | x_i \epsilon Z/\{0\}, i = 1, 2, \ldots, n \text{ and } x_i \neq x_j, \text{ if } i \neq j\}.$$

Two subsets of Ω_n of interest are

$$\Omega_n^+ = \{(0, x_1, \ldots, x_n) | (0, x_1, \ldots, x_n) \epsilon \Omega_n \text{ and } x_i > 0 \text{ for } i = 1, 2, \ldots, n\}$$

and

$$\Omega_n^- = \{(0, x_1, \ldots, x_n) | (0, x_1, \ldots, x_n) \epsilon \Omega_n \text{ and } x_i < 0 \text{ for } i = 1, 2, \ldots, n\}.$$

Elements of these sets will be denoted by ω_n. Ω_n contains all self-avoiding paths of n steps starting at the origin, while Ω_n^+ and Ω_n^- contain all such paths which exist to one side or the other of the origin.

If $\omega_n \epsilon \Omega_n$ and $\omega_m \epsilon \Omega_m$ where $\omega_n = (0, x_1, \ldots, x_n)$ and $\omega_m = (0, y_1, \ldots, y_m)$ then define the <u>concatenation</u> of ω_n with ω_m (denoted $\omega_n * \omega_m$) by

$$\omega_n * \omega_m = (0, x_1, \ldots, x_n, x_n + y_1, \ldots, x_n + y_m).$$

It is not necessarily true that $\omega_n * \omega_m \epsilon \Omega_{n+m}$, although it would be if, for example, $x_n = \max\{0, x_1, \ldots, x_n\}$ and $x_m \epsilon \Omega_m^+$.

Returning to our i.i.d. sequence we note that the sequence of partial sums $\{S_n\}$ may be regarded as a random walk on Z. Given the assumptions on $G(\cdot)$, the random walk will step eventually to a site previously occupied, or alternatively, will try to fill a point already filled. This occurs at time

$$T = \min[n \geq 1 | S_n \epsilon \{0, S_1, \ldots, S_{n-1}\}].$$

If the random walk is stopped at time T, then at time n, the random walk is still proceeding on the set

$$A_n = \{T > n\},$$

which we will refer to as the set of non-extinction. As defined on page 105 of [7], the self-avoiding random walk is the sequence $0, S_1, \ldots, S_n$ conditioned on A_n.

In order to study growth properties of $S_0, \ldots, S_n$ on A_n, it is necessary to study the distribution of self-avoiding paths of n steps. This distribution is given by

$$p_n(\omega_n) = \frac{P[(0, S_1, \ldots, S_n) = \omega_n]}{\sum\limits_{\omega'_n \in \Omega_n} P[(0, S_1, \ldots, S_n) = \omega'_n]}.$$

Several problems arise in trying this, not the least of which is that the $p_n(\cdot)$'s are not consistent because the denominator does not factor. That is to say the distribution of $S_0, \ldots, S_n$ conditioned on A_n is not necessarily the distribution of $S_0, \ldots, S_n$ conditioned on A_{n+m}. A natural approach is to look for a well behaved subsequence where one can not only get consistent finite dimensional probabilities, but also factor the $p_n(\cdot)$'s. This subsequence is defined by the "break points" of the next section. These break points allow the process to be divided into independent pieces. With further assumptions made on $G(\cdot)$, one can then get conditional limit results for S_n. Since we will be looking at the process on a subsequence, we define (for a nonnegative, integer-valued random variable T_1)

$$A_{T_1} = \{T_1 < T\},$$

$$A^+_{T_1} = A_{T_1} \cap \{S_{T_1} > 0\},$$

and

$$A^-_{T_1} = A_{T_1} \cap \{S_{T_1} < 0\}.$$

3. Break Points

Since the main problem in studying the process arises from the complex stochastic conditioning, one would like to avoid the necessity of conditioning on the entire past. The idea here is to find time points where the future of the process haves independently of the past. Intuitively, it is clear that the self-avoiding walk will cross zero finitely often and then remain on one side, especially in the case where the random variables are bounded. So, eventually the process will not crossover certain points again. However, if such a point of no crossover were a relative maxima or minima, then the future would not depend on the configuration of "filled" points on the other side of these points of no crossover. Define these relative maxima or minima where there is no future crossover before extinction to be break points, since they break the future and past into independent pieces. Let

$$M_1 = \inf_{k\geq 1}\{S_k \text{ is a relative maxima and } S_{k+m} \geq S_k \text{ if } k+m \leq T \text{ for } m \geq 1\}$$

$$\vdots$$

$$M_{n+1} = \inf_{k>M_1+\ldots+M_n}\{S_k \text{ is a relative maxima and } S_{k+m} \geq S_k \text{ if } k+m \leq T \text{ for } m \geq 1\}$$

and

$$N_1 = \inf_{k\geq 1}\{S_k \text{ is a relative minima and } S_{k+m} \geq S_k \text{ if } k+m \leq T \text{ for } m \geq 1\}$$

$$\vdots$$

$$N_{n+1} = \inf_{k>N_1+\ldots+N_n}\{S_k \text{ is a relative minima and } S_{k+m} \geq S_k \text{ if } k+m \leq T \text{ for } m \geq 1\}$$

(We take the infimum of the empty set to be ∞.)

Now, for $i = 1, 2, \ldots$, define

$$T_i = \min[M_i, N_i].$$

As a notational convenience, define for $n = 1, 2, \ldots$,

$$\tau_n = \sum_{i=1}^{n} T_i.$$

Also let

$$\omega_{T_1} = (0, S_1, \ldots, S_{\tau_1}) \qquad \text{on } \{\tau_1 < T\}$$

$$\omega_{T_2} = (0, S_{\tau_1+1} - S_{\tau_1}, \ldots, S_{\tau_2} - S_{\tau_1}) \qquad \text{on } \{\tau_2 < T\}$$

$$\vdots$$

$$\omega_{T_{n+1}} = (0, S_{\tau_n+1} - S_{\tau_n}, \ldots, S_{\tau_{n+1}} - S_{\tau_n}) \qquad \text{on } \{\tau_{n+1} < T\}.$$

Clearly, on the appropriate set of nonextinction,

$$\omega_{T_1} \epsilon \Omega$$

and

$$\omega_{T_i} \epsilon \Omega^+ \cup \Omega^-$$

for $i = 2, 3, \ldots$, where

$$\Omega = \cup_n \Omega_n, \qquad \Omega^+ = \cup_n \Omega_n^+, \qquad \Omega^- = \cup_n \Omega_n^-.$$

Note that conditioned on the set A_{τ_n}

$$(0, S_1, \ldots, S_{\tau_n}) = \omega_{T_1} * \omega_{T_2} * \ldots * \omega_{T_n}.$$

To examine distributional properties of the concatenation of these strings of "random fields", let

$$p_{T_j}^+(\omega_n) = P(\omega_{T_j} = \omega_n | A_{\tau_j}^+, \omega_n \epsilon \Omega_n)$$

and

$$p_{T_j}^-(\omega_n) = P(\omega_n | A_{\tau_j}^-, \omega_n \epsilon \Omega_n).$$

The distribution $p_{T_1}^+(\cdot)$ is different than $p_{T_j}^+(\cdot)$, for $j \geq 2$ since steps below zero are permitted before T_1 on $\{T_1 < \infty\}$, while $p_{T_j}^+(\omega_n) = 0, j \geq 2$, if ω_n contains a step below zero. Note that as a consequence, for all $j \geq 2, p_{T_j}^+(\cdot)$ has all of its mass on Ω^+, while $p_{T_j}^-(\cdot)$ has all of its mass on Ω^-.

If each term of the sequence $\{\tau_n\}, n = 1, 2, \ldots$, were a stopping time for the sequence $\{X_i\}$, then by the strong Markov property the sequence $X_{\tau_n+1}, X_{\tau_n+2}, \ldots$, would be independent of the past and distributed as $X_1, X_2, \ldots$. This would permit the factorization of $P(0, S_1, \ldots, S_{\tau_n} | A_{\tau_n})$. It is somewhat surprising that while $\tau_n, n \geq 1$, is never a stopping time, something like the strong Markov property still holds; so that $P(0, S_1, \ldots, S_{\tau_n} | A_{\tau_n}^+)$ can still be factored as shown in the next proposition. It is also shown that $p_{T_2}^+(\omega_n) = p_{T_j}^+(\omega_n), j \geq 3$.

Proposition 1 *Assume that $G(\cdot)$ is integer valued with no mass at 0.*

$$P(0, S_1, \ldots, S_{\tau_k} = \omega_{n_1} * \omega_{n_2} * \ldots * \omega_{n_k} | A_{\tau_k}^+) = p_{T_1}^+(\omega_{n_1}) \cdot \prod_{j=2}^{k} p_{T_j}^+(\omega_{n_j}).$$

Proof: It will be shown for each n, that conditioned on the set $A^+_{\tau_k}$, the distribution of $(0, S_{\tau_k+1} - S_{\tau_k}, \ldots, S_T - S_{\tau_k})$ is

1. exactly the same as the distributrion of $(0, S_1, \ldots, S_T)$ conditioned on the set $\{S_i \geq 0, i = 0, 1, \ldots, T\}$, and
2. is independent of $(0, S_1, \ldots S_{\tau_k})$.

Let $U_1, U_2, \ldots$ denote times of occurrence of successive relative maxima of the sequence $\{S_n\}$. Note that for any n, $P(\tau_k = U_m$ for some $m \geq k | A^+_{\tau_k}) = 1$, i.e. any break point on the positive side of the axis is, at some time point, a relative maxima. Note also that for every $M \geq 1, U_m$ is a stopping time for the sequence $\{X_i\}$, for by the strong Markov property the sequence $(X_{U_m+1}, X_{U_m+2}, \ldots)$ is distributed exactly the same as $(X_1, X_2, \ldots)$ and is independent of $(X_1, X_2, \ldots, X_{U_m})$. As a consequence, $(0, S_{U_m+1} - S_{U_m}, S_{U_m+2} - S_{U_m}, \ldots)$ is distributed as $(0, S_1, S_2, \ldots)$ and is independent of $(0, S_1, S_2, \ldots, S_{U_m})$. This holds in the particular case where $(), S_1, S_2, \ldots, S_{U_m})$ is self-avoiding, i.e., on the set $\{U_m < T\}$. On this set, S_{U_m} is a break point if and only if $\{S_{U_m+i} - S_{U_m} \geq 0, i = 1, 2, \ldots, T\}$ holds, i.e., the sequence has the same distribution as $(0, S_1, \ldots, S_T)$ on the set $\{S_i \geq 0, i = 1, 2, \ldots, T\}$, and is independent of $(0, S_1, \ldots, S_{U_m})$ which is equal to $(0, S_1, \ldots, S_{\tau_k})$ for some k. □

4. Limit Behavior

Since the break points divide the point process into independent pieces, the problem of conditioning on the past is greatly reduced on the subsequence S_{τ_n}. This being the case, one can then study the growth on this subsequence using the results on conditional independence of the last section. Define

$$p_{nk} = P[(0, S_1, \ldots, S_n) \text{ has at least } k \text{ points of no crossover } | A^+_n). \tag{1}$$

The growth is linear in the following sense.

Theorem 2 *Let $G(\cdot)$ satisfy the following conditions:*

1. *$G(\cdot)$ is integer valued with no mass at 0.*

2. *$E_G(|X|) < \infty$.*

3. *For every integer $k, \lim_{n\to\infty} p_{nk} = 1$.*

Then there exists a positive constant $m > 0$, depending on $G(\cdot)$, such that

$$P\left(\left|\frac{S_n}{n} - m\right| > \epsilon | A_n^+\right) \longrightarrow 0$$

as $n \to \infty$.

Proof: In order to prove the result for the self-avoiding random walk, we first prove it along the subsequence of break points. We state this preliminary result as a proposition which we will prove later.

Proposition 3 *Under the conditions of the theorem,*

$$P\left(\left|\frac{S_{\tau_n}}{\tau_n} - m\right| > \epsilon | A_{\tau_n}^+\right) \longrightarrow 0$$

as $n \to \infty$.

Assuming this result for a (random) subsequence of the self-avoiding random walk, we can get the result for the walk itself, when $\tau_k < n < \tau_{k+1} < T$, by noting the inequality:

$$S_{\tau_k} \le S_n \le S_{\tau_k+1} \qquad \text{on } A_{\tau_{k+1}}^+ \text{ for } \tau_k \le n \le \tau_{k+1}.$$

Thus on $A_{\tau_{k+1}}^+$

$$\frac{S_{\tau_k}}{\tau_{k+1}} \le \frac{S_n}{n} \le \frac{S_{\tau_{k+1}}}{\tau_k}.$$

From Proposition 3, we know

$$\frac{S_{\tau_k}}{\tau_k} \longrightarrow m \qquad \text{in probability,}$$

and, since the times between break points form a renewal sequence,

$$\frac{\tau_k}{\tau_{k+1}} \longrightarrow 1 \qquad \text{in probability.}$$

Thus

$$\frac{S_n}{n} \longrightarrow m \qquad \text{in probability.}$$

However, this doesn't work when

$$\tau_k \leq n \leq T$$

and τ_k is the last break point before extinction. The following lemma provides us with the necessary result from renewal theory to complete our proof.

Unfortunately we must first introduce some more notation. Let $(V_1, W_1), (V_2, W_2), \ldots$ be a sequence of i.i.d. random vectors defined on $\mathcal{N} \times \mathcal{N}$, where $\mathcal{N}$ is the natural numbers. Let $P(\cdot, \cdot)$ be their probability distribution, and assume $P(k, k) = 0$ for all $k \epsilon \mathcal{N}$. Define

$$\begin{aligned}
T &= \int \left\{ \left(\sum_{i=1}^{k-1} X_i \right) + Y_k : Y_k < X_k \right\} \\
N(n) &= \sup \left\{ k : \sum_{i=1}^{k} X_i \leq n \right\} \\
\eta_n &= n - \sum_{i=1}^{N(n)} X_i \qquad \text{if } n < T \\
\zeta_n &= \min \left\{ T - n, \sum_{i=1}^{N(n)+1} X_i - n \right\} \qquad \text{if } n < T \\
A(n) &= \begin{cases} 1 & \text{if } n < T \\ 0 & \text{otherwise} \end{cases} \\
Q(i) &= \sum_{k_2 = i+1}^{\infty} P(i, k_2).
\end{aligned}$$

Lemma 4 *If there exists λ such that $\Sigma_{i=1}^{\infty} e^{\lambda i} Q(i) = 1$, and $\Sigma_{i=1}^{\infty} i e^{\lambda i} Q(i) = \mu < \infty$, then*

$$\lim_{n \to \infty} P[\eta_n \leq x | A(n) = 1] = \frac{\Sigma_{i=1}^{x} P[\min(X, Y) > i] e^{\lambda i}}{\Sigma_{i=1}^{\infty} P[\min(X, Y) > i] e^{\lambda i}}$$

and

$$\lim_{n \to \infty} P[\zeta_n \leq x | A(n) = 1] = \frac{\Sigma_{i=1}^{\infty} P[i < \min(X, Y) < i + x] e^{\lambda i}}{\Sigma_{i=1}^{\infty} P[\min(X, Y) > i] e^{\lambda i}}.$$

Now, returning to the proof of the theorem, we can write

$$S_n = S_{\tau_{k_n}} + \sum_{i=1}^{\eta_n} X_i,$$

where τ_{k_n} is the last break point before $n, \eta_n = n - \tau_{k_n}$, and $X_1, \ldots, X_{\eta_n}$ are the step sizes. From the lemma we see that η_n has a finite limiting distribution, and on the set A_n^+ we know that for $k > 0, S_{\tau_k}$ and $\Sigma_{i=1}^{\eta_n} X_i$ are both positive. Thus dividing by n, we get

$$\frac{S_n}{n} = \frac{S_{\tau_{k_n}} + \Sigma_{i=1}^{\eta_n} X_i}{\tau_{k_n} + \eta_n} = \frac{S_{\tau_{k_n}}}{\tau_{k_n} + \eta_n} + \frac{\sum\limits_{i=1}^{\eta_n} X_i}{\tau_{k_n} + \eta_n}.$$

The second term converges to 0 in probability as $n \to \infty$, and the first term converges to m in probability since

$$\frac{\tau_{k_n} + \eta_n}{\tau_{k_n}} \longrightarrow 1$$

in probability as $n \to \infty$. This completes the proof of the theorem. □

We devote the rest of this section to the proofs of Proposition 3 and Lemma 4.

Proof of Proposition 3: Owing to the i.i.d. nature of the ω_T, for $i \geq 2$ from Proposition 1, it suffices to show that

$$E(|S_{T_1+T_2} - S_{T_1}||A^+_{T_1+T_2}) = c_1 < \infty$$

and

$$E(T_2|A^+_{T_1+T_2}) = c_2 < \infty,$$

in which case we may set $m = c_1/c_2$ and be done.

Consider $P(T_1 + T_2 \geq n|T_1 + T_2 < T)$. Clearly $P(T \geq n) \leq k \cdot p^n$, where $k > 0$ and $p < 1$, i.e. the time to extinction is a random variable, the tail of whose distribution decreases geometrically with time. Evidently

$$P(T_1 + T_2 \geq n|T_1 + T_2 < T) = P(T > T_1 + T_2 \geq n)/P(T_1 + T_2 < T) \leq k'p^n,$$

where

$$k' = k/P(T_1 + T_2 < T);$$

so that the conditional moments at $T_1, T_2, \ldots$ all exist and are finite. Since the length of a step has finite expectation, the conditional expectation of $S_{\tau_{n+1}} - S_{\tau_n}$ also exists and is finite. Thus

$$P(|S_{\tau_n}/n - c_1| > \epsilon | A^+_{\tau_n}) \to 0$$

as well as

$$P(|\tau_n/n - c_2| > \epsilon | A_{\tau_n}) \to 0,$$

yielding the result. □

Proof of Lemma 4: By conditioning on the value of (V_1, W_1) we obtain the following:

$$P[A(n) = 1] = P[\min(V_1, W_1) > n] + \sum_{i=1}^{n} P[A(n-i) = 1]Q(i),$$

$$P[\eta_n \leq x, A(n) = 1] = P[\min(V_1, W_1) > n]\chi_{[0,x]}(n) + \sum_{i=1}^{n} P[\eta_{n-i} \leq x, A(n-i) = 1]Q(i),$$

and

$$P[\zeta_n \leq x, A(n) = 1] = P[n < \min(V_1, W_1) < n + x] + \sum_{i=1}^{n} P[\zeta_{n-i} \leq x, A(n-i) = 1]Q(i).$$

By the assumptions of the lemma, we can apply the renewal theorem to $e^{\lambda n}P[A(n) = 1], e^{\lambda n}P[\eta_n \leq x, A(n) = 1]$, and $e^{\lambda n}P[\zeta_n \leq x, A(n) = 1]$ to get

$$\lim_{n\to\infty} e^{\lambda n}P[A(n) = 1] = \frac{1}{\mu}\sum_{i=1}^{\infty} P[\min(X, Y) > i]e^{\lambda i},$$

$$\lim_{n\to\infty} e^{\lambda n}P[\eta_n \leq x, A(n) = 1] = \frac{1}{\mu}\sum_{i=1}^{x} P[\min(X, Y) > i]e^{\lambda i},$$

$$\lim_{n\to\infty} e^{\lambda n}P[\zeta_n \leq x, A(n) = 1] = \frac{1}{\mu}\sum_{i=1}^{\infty} P[i < \min(X, Y) < i + x]e^{\lambda i}.$$

Division now gives us the desired results. □

5. Some Further Results

In proving our results on the growth rate of the self-avoiding random walk, we have made the assumption that break points occur quite often, at least if the process itself lasts for a long period of time. In this section we would like to show that at least for some distributions $G(\cdot)$, break points are a common feature of the sequence $(0, S_1, \ldots, S_n)$ on the set A_n for large n. It is probably true under very general assumptions on $G(\cdot)$, but difficult to prove owing to the complicated conditioning involved. As will be seen in the following proposition, if $G(\cdot)$ satisfies assumption A, then the conditional probability that the sequence contains at least k points of no crossover can be made arbitrarily high by letting $n \to \infty$. Let p_{nk} be as defined in equation 1.

Proposition 5 *If $G(\cdot)$ has all of its mass on a bounded interval $[-r, r]$ and $P(X = i) > 0$ for every integer i for which $-r \leq i \leq r$, then for every integer $k, \lim_{n\to\infty} p_{nk} = 1$.*

Proof: Fix $M > 0, \epsilon > 0$, and let n_M be such that for all $n \geq n_M, P(|S_n| > M|A_n^+) > 1 - \epsilon$. If one can show that conditioned on the set $\{|S_n| > M\} \cap A_n$, that k points of no crossover exist with arbitrarily high probability, then the proof will be complete. On this set, suppose, without loss of generality, that $S_n > 0$. By assumption A, steps are $\leq r$ in length. Suppose that $M = 6 \cdot r \cdot j$ for some positive integer j. Thinking of the sequence $(0, S_1, \ldots, S_n)$ as just being a collection of filled points, select a subsequence $M_1, M_2, \ldots, M_j$ such that $0 < M_1 < \ldots < M_j$ and $M_{i+1} - M_i \geq 2 \cdot r + 1$ for $i = 1, 2, \ldots, j - 1$. (Such a subsequence must always exist, e.g. M_i = largest filled lattice point in $[3r \cdot i, 3r \cdot i + r]$.) Suppose that $\omega_n = (0, S_1, \ldots, S_n)$ is such that M_i is a point of crossover. It will be shown that there exists an element ω_n' of Ω_n such that ω_n and ω_n' take the same steps off of $[M_i - r, M_i + r]$, and $p_n(\omega_n') \geq c \cdot p_n(\omega_n)$, where $c = [(\min_{i=1,\ldots,r} P(X_1 = i))/\max_{i=1,\ldots,r} P(X_1 = i)]^{2\cdot r}$. Whether or not M_i is a point of crossover is determined by steps from lattice points in $[M_i - r, M_i + r]$ to points in the same interval, since steps are of length r or less. There are at most $2 \cdot r$ such steps. If M_i is a point of crossover for ω_n, it is necessary to show that a rearrangement of these at most $2 \cdot r$ steps exists, yielding an ω_n' where M_i is a point of no crossover. For simplicity, suppose that $G(\cdot)$ is such that a point can only be crossed over twice, i.e. $r = 3$. The argument for this case easily generalizes.

Suppose that the path ω_n is such that the self-avoiding walk does cross over M_i. Then there are points q_1, q_2, q_3 to the left of M_i, and points p_1, p_2, p_3 to the right of M_i, such that the following hold. From q_1 we step to M_i in one step, from M_i we step to p_1 in one step, and from p_1 we wander among points to the right of M_i until we step to p_2; from p_2 we crossover M_i and hit q_2; from q_2 we wander among points to the left of M_i until stepping to q_3, from which we crossover M_i in one step to p_3, whereupon we remain to the right of M_i. Thus these points comprise the places where M_i is crossed over, or which step to M_i, or which M_i steps to. This sequence of steps,

$$q_1 \longrightarrow M_i \longrightarrow p_1 \longrightarrow \ldots \longrightarrow p_2 \longrightarrow q_2 \longrightarrow \ldots \longrightarrow q_3 \longrightarrow p_3,$$

where the arrows indicate a single step, is displayed in figure 1. Conditioning on the location of $q_1, q_2, q_3, p_1, p_2, p_3$ and leaving segments $p_1 \to \ldots \to p_2$ and $q_2 \to \ldots \to q_3$ undisturbed, we change a few steps obtaining a new sequence ω'_n,

$$q_1 \longrightarrow q_2 \longrightarrow \ldots \longrightarrow q_3 \longrightarrow M_i \longrightarrow p_1 \longrightarrow \ldots \longrightarrow p_2 \longrightarrow p_3,$$

which is displayed in figure 2. Thus for every path ω_n where M_i is a point of crossover, by changing at most $2r$ steps in $[M_i - r, M_i + r]$, a path ω'_n is obtained for which M_i is a point of no crossover. In addition, since at most $2r$ steps were changed, $p_n(\omega'_n) \geq c \cdot p_n(\omega_n)$. Given this collection of filled points, the probability that any particular M_i is a point of no crossover is at least $c/(1+c)$. By the method of construction of ω'_n, this probability is independent of the behavior of ω_n at $M_1, \ldots, M_{i-1}$. Therefore, the distribution of the number of points of no crossover among $M_1, \ldots, M_j$ is bounded from below by a binomial $(j, c/(1+c))$ random variable, that is

$$p_{nk} \geq \sum_{i=k}^{j} \binom{j}{i} \left(\frac{c}{1+c}\right)^i \left(\frac{1}{1+c}\right)^{j-i}.$$

As M (or equivalently j) is made arbitrarily large, p_{nk} can be made arbitrarily close to one. □

6. Discussion

We have applied a technique for finding regeneration points, termed "break points", to a self-avoiding random walk in order to study its limit behavior. While these regeneration points do not occur at stopping times, they still allow the decomposition of the process into independent pieces which greatly simplified the study of its limit behavior. The technique was also applicable to an entirely different problem in percolation which suggests even wider applicability. The key to the technique is asking the correct question about future behavior which will yield the regeneration behavior.

References

Kuczek, T. (1989): The central limit theorem for the right edge of supercritical oriented percolation. *Ann. Prob.*, **17**, No. 4, 1322–1332.

Murphree, E.S. (1987): The distribution of delays in transient renewal processes. *Sankhyā: The Indian Journal of Statistics*, **49**, Series A, Part 1, 113–121.

Spitzer, F. (1954): *Principles of Random Walk*, Princeton: Wiley.