Elect. Comm. in Probab. **8** (2003) 135–141

ELECTRONIC COMMUNICATIONS in PROBABILITY

STRICT CONVEXITY OF THE LIMIT SHAPE IN FIRST-PASSAGE PERCOLATION

STEVEN P. LALLEY[1]
Department of Statistics, University of Chicago, Chicago IL 60637 USA
email: `lalley@galton.uchicago.edu`

Submitted 26 September 2003, *accepted in final form* 8 October 2003

AMS 2000 Subject classification: 60K35
Keywords: Shape theorem, first-passage percolation, oriented percolation

Abstract
Sufficient conditions are given for the strict convexity of the limit shape in standard first-passage percolation. These conditions involve (1) asymptotic "straightness" of the geodesics, and (2) existence of mean-zero limit distributions for the first-passage times.

1 Introduction

The limit shape of the infected set in standard first-passage percolation (see [5] and [7] for an introduction to the subject) is well known to be convex. It is widely believed that, with few exceptions (perhaps only when the distribution F of the passage time of an edge is supported by two points), the limit set $\mathcal{B}$ is *strictly* convex in the sense that line segments connecting any two points of $\mathcal{B}$ are entirely contained, except for their endpoints, in the interior of $\mathcal{B}$. Strict convexity of the limit set plays an important role in certain questions regarding the evolution of multi-type stochastic competition models – see, for instance, [8]. Nevertheless, strict convexity has not yet been established for any distribution F, not even the exponential distribution.

It has been observed by Durrett and Liggett [2] that when the edge passage time distribution F is supported by two points, the limit shape need not be strictly convex: its boundary may include nontrivial sections of hyperplanes. This situation seems to be atypical in at least two other important respects. First, time-minimizing paths from the origin to distant points need not be (to first order) straight. Second, the first-passage distributions are, for distant points in directions along the "flat spots", highly concentrated: In particular, if $T(x)$ is the first passage to the point of $\mathbb{Z}^d$ nearest to $x \in \mathbb{R}^d$ and a is the smaller point in the support of F, then for certain directions u, as $n \to \infty$,

$$T(nu) - na \xrightarrow{\mathcal{D}} G \tag{1}$$

for a certain distribution G independent of the direction u.

[1] RESEARCH SUPPORTED BY NSF GRANT DMS-0071970

That the latter two properties are indeed closely related to the strict convexity of the the limit shape is further suggested by more recent results of Newman, Piza, and Licea [13], [12], [10], [9]. They have shown that, under certain additional hypotheses on F, if the limit shape is not only strictly convex but *uniformly curved* then (a) the variance of $T(nu)$ must grow at least logarithmically in n (see also [14] for another proof of this in the case where F is an exponential distribution), and (b) geodesic segments (time-minimizing paths) between distant points of $\mathbb{Z}^d$ must be (to first order) straight line segments.
The purpose of this note is to show that properties closely related to (a) and (b) above are in fact *sufficient* for the strict convexity of the limit shape $\mathcal{B}$. Assume that the passage times for different edges of the integer lattice $\mathbb{Z}^d$ are independent and identically distributed, and that their common distribution F is supported by $[0, \infty)$, is nonatomic, and has finite exponential moment $\int e^{\gamma t}\, dF(t)$ for some $\gamma > 0$. These assumptions ensure that time-minimizing paths between points x, y of $\mathbb{Z}^d$ are unique, and that the Shape Theorem holds, with a compact, convex limit shape $\mathcal{B}$ containing the origin in its interior — see Theorems 1.7 and 1.15 of [7]. For each $x \in \mathbb{R}^d \setminus \{0\}$ define

$$\mu(x) = \lim_{n\to\infty} \frac{ET(nx)}{n} \tag{2}$$

to be the inverse infection speed in direction x. The limit shape $\mathcal{B}$ consists of the origin and those $x \neq 0$ such that $\mu(x) \leq 1$.
Let u be a fixed nonzero vector in $\mathbb{R}^d$, and let $\mathcal{L}_u$ be the ray through u emanating from the origin.

Hypothesis 1 *For any convex cone $\mathcal{A}$ of $\mathbb{R}^d$ containing the vector u in its interior, and for each $\delta > 0$ there exists $R = R(\delta, \mathcal{A}) < \infty$ such that the following is true: For each point $v \in \mathbb{Z}^d \cap \mathcal{A}$ at distance ≤ 2 from the line $\mathcal{L}_u$, the probability that the time-minimizing path from the origin to v is contained in $\mathcal{A} \cup \{x : \|x\| \leq R\}$ is at least $1 - \delta$.*

In the following, the only convex cones considered will be those whose intersections with the unit sphere $\Sigma_{d-1} \subset \mathbb{R}^d$ are spherical caps centered at the point $\mathcal{L}_u \cap \Sigma_{d-1}$; such cones will be called *spherical cones.* Two spherical cones will be said to have the same *aperture* if the spherical caps that determine them are congruent.

Hypothesis 2 *There exist a mean-zero probability distribution G_u n the real line and a scalar sequence $a(n) \to \infty$ such that as $n \to \infty$,*

$$\frac{T(nu) - n\mu(u)}{a(n)} \xrightarrow{\mathcal{D}} G_u. \tag{3}$$

Theorem 1 *Let u, v be linearly independent vectors in $\mathbb{R}^d$. If Hypotheses 1 – 2 hold for both u and v then for each $\alpha \in (0, 1)$,*

$$\mu(\alpha u + (1-\alpha)v) < \alpha\mu(u) + (1-\alpha)\mu(v). \tag{4}$$

Theorem 1 clearly implies that if the hypotheses 1 and 2 hold for all u in a dense set of unit vectors then the limit shape $\mathcal{B}$ is strictly convex. Unfortunately, the convergence in law of the rescaled first-passage times $T(nu)$ has not been established for any distribution F, nor is it at all clear that even if such convergence holds the limit distribution G_u should have mean zero. It has been conjectured that the variance of $T(nu)$ is of order $n^{2/3}$ at least for continuous distributions F with finite exponential moments, and it is now known [1] that at least for certain distributions the variance of $T(nu)$ grows sublinearly with n. This suggests that a limiting distribution, if it exists, may not be Gaussian, but rather a Tracy-Widom distribution. See Johansson [6] for results in this direction on what seems to be a closely related model.

2 Proof of Theorem 1

The proof, like that of [2], makes use of a derived oriented percolation process. This oriented percolation is always 2−dimensional, regardless of the ambient dimension of the first-passage percolation, and, in the general case, is a combined edge+site percolation: Bernoulli-p_v random variables are attached to the *vertical* edges, Bernoulli-p_h random variables are attached to the *horizontal* edges, and Bernoulli-p_s random variables are attached to the vertices of the lattice $\mathbb{Z}^2$; these random variables are mutually independent. An *open path* in the oriented percolation is a finite alternating sequence e_i, v_i of edges and vertices, all of which are "open" (that is, the attached Bernoulli random variables all take the value 1), with each vertex v_i incident to the edges e_i and e_{i+1}, and such that the sequence of vertices visited by successive edges in the path is nondecreasing with respect to the usual partial order on $\mathbb{Z}^d$. If there is an infinite open path starting at the origin, then *percolation* is said to occur. When percolation occurs, an infinite open path starting at the origin may visit the diagonal infinitely often; denote this event by Ω.

Proposition 3 *There is a critical value $p_c < 1$ such that if p_v, p_h and p_s all exceed p_c then $P(\Omega) > 0$.*

See [3] for the proof in the case where $p_v = p_h$ and $p_s = 1$; the extension stated here may be proved in much the same manner.

Assume now that the hypotheses of Theorem 1 are in force. For any path γ in the integer lattice, let $\tau(\gamma)$ denote the time required to traverse γ (that is, the sum of the edge passage times over all edges in γ). The derived oriented percolation process on $\mathbb{Z}^2$ will be defined by making an appropriate embedding of $\mathbb{Z}^2$ into the ambient space $\mathbb{Z}^d$ and using the travel times $\tau(\gamma)$ along appropriate paths to determine which edges and vertices of $\mathbb{Z}^2$ will be open. The embedding $\varphi : \mathbb{Z}^2 \to \mathbb{Z}^d$ will be a linear mapping that sends $(1, 0)$ to x and $(0, 1)$ to y for certain linearly independent vectors $x, y \in \mathbb{Z}^d$. Along with this embedding there is a natural associated mapping, also denoted by φ, of the edges of the lattice $\mathbb{Z}^2$ to line segments in $\mathbb{R}^d$. The mapping φ of vertices and edges will be augmented by the following geometric objects: (A) Each embedded vertex $\varphi(v)$ will be surrounded by a closed ball B_v of radius R centered at $\varphi(v)$; the constant R will be the same for all vertices, and small enough that no two distinct balls $B_v, B_{v'}$ intersect. (B) Each embedded edge $\varphi(e)$ will be enclosed in a region D_e that consists of the intersection of two congruent spherical cones, one based at each endpoint of $\varphi(v)$. The cones used in this construction will all have the same aperture, and the aperture will be sufficiently small that no two of the regions $D_e, D_{e'}$ overlap except possibly at the vertices $\varphi(v)$. See Figure 1 for a depiction of a portion of the decorated lattice.

Bernoulli random variables X_e, X_v indexed by the edges e and vertices v are now defined as follows: For each edge e of $\mathbb{Z}^2$, denote by $\psi(e)$ the segment of the edge $\varphi(e)$ exterior to the balls B_v and $B_{v'}$, where v, v' are the vertices incident to e, and let $z = z_e, z' = z'_e$ be the lattice points closest to the endpoints of the segment $\psi(e)$. Define Y_e to be the minimum traversal

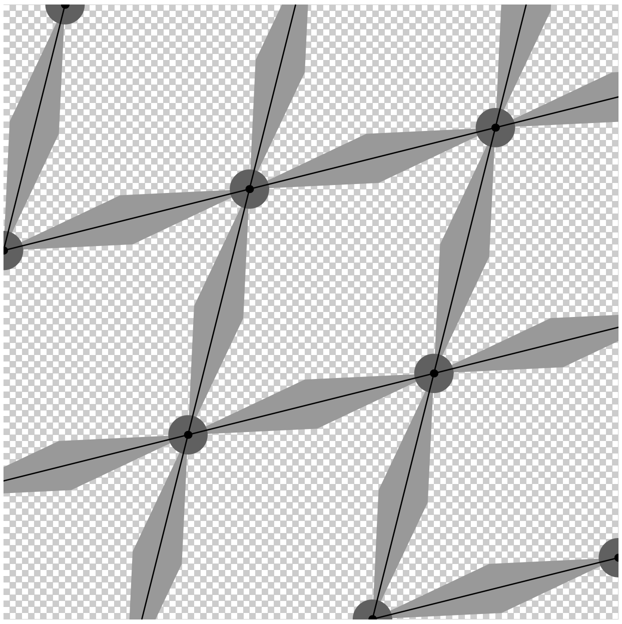

Figure 1:

time $\tau(\gamma)$ among all paths γ from z to z' that lie entirely in the region D_e. Set

$$X_e = 1 \quad \text{if} \quad Y_e \leq t_H \quad \text{and } e \text{ is horizontal;} \tag{5}$$

$$X_e = 1 \quad \text{if} \quad Y_e \leq t_V \quad \text{and } e \text{ is vertical;} \tag{6}$$

and $X_e = 0$ otherwise. The constants t_H, t_V will be specified later. Observe that because the regions D_e are nonoverlapping, the random variables Y_e, and hence also X_e, are mutually independent. For each vertex $v \in \mathbb{Z}^2$, set

$$X_v = 1 \tag{7}$$

if the passage times of all edges in the ball B_v are bounded above by t_S, a constant to be specified later; otherwise, $X_v = 0$.

Note that if the edge passage time distribution F has support contained in $[0, t_S]$ then $X_v = 1$ almost surely for each vertex v; in this case, the percolation event Ω is determined completely by the edge r.v.s X_e. In general, if the support of F is unbounded, then it will not be the case that $X_v = 1$ a.s.; moreover, the random variables X_e and X_v will not necessarily be mutually independent, and so Proposition 3 will not apply directly. A modification of the construction to remedy this difficulty will be proposed later. However, the random variables X_e indexed by edges e will be mutually independent, and $P\{X_e = 1\}$ will depend only on whether the edge e is vertical or horizontal: denote the two possible values by p_V, p_H, respectively. Similarly, the distribution of the travel time Y_e depends only on whether e is vertical or horizontal.

To complete the construction, it remains to specify the constants R, t_H, t_V, t_S, the vectors x, y, and the aperture of the spherical cone that determines the regions D_e. Here Hypotheses 1 – 2 come into play: Hypothesis 1 will be used to determine R and the aperture of the cone, and Hypothesis 2 the remaining parameters. Let u, v be linearly independent vectors in $\mathbb{R}^d$ for which Hypothesis 1 holds. Given this choice u, v, choose an aperture sufficiently small that

$$\mathcal{A}_u \cap \mathcal{A}_v = \{0\}, \tag{8}$$

where $\mathcal{A}_u, \mathcal{A}_v$ are the closed spherical cones of this aperture centered at the points where the rays $\mathcal{L}_u, \mathcal{L}_v$ intersect the unit sphere Σ_{d-1}. By Hypothesis 1, for any $\delta > 0$ there exists $1 < R < \infty$ such that for any points $\xi, \zeta \in \mathbb{Z}^d$ both of norm at least $R/2$ and both at distance less than 2 from the ray $\mathcal{L}_u$ (alternatively, the ray $\mathcal{L}_v$), there is probability at least $1-\delta$ that the geodesic connecting the points ξ, ζ lie entirely in the cone $\mathcal{A}_u$ (alternatively, the cone $\mathcal{A}_v$). Consequently, if R is used as the radius of the balls B_v and if each region D_e is the intersection of two spherical cones each congruent to $\mathcal{A}_u$ (and therefore also $\mathcal{A}_v$), then for each edge e the probability that the geodesic segment connecting the lattice points z_e and z'_e lies entirely in D_e is at least $1-2\delta$.
Assume now that the vectors u, v satisfy both Hypotheses 1 and 2. Let R and the cone aperture be chosen as in the preceding paragraph (these aso depend on $\delta > 0$, which will be specified below). According to Hypothesis 2, there are scalar sequences $a_u(n), a_v(n)$ so that the travel times $T(nu), T(nv)$, normalized as in (3), have mean-zero limit distributions G_u, G_v. Because these limit distributions have mean zero, if g_u, g_v are $(1-\epsilon)$th quantiles of the distributions G_u, G_v, for any $\epsilon > 0$, then

$$\nu_u := \frac{\int_{-\infty}^{g_u} t\, dG_u(t)}{1 - G_u(g_u)} < 0 \qquad \text{and} \qquad \nu_v := \frac{\int_{-\infty}^{g_v} t\, dG_v(t)}{1 - G_v(g_v)} < 0.$$

This is in fact the crucial point of the proof. The plan now is to let x, y be the points of $\mathbb{Z}^d$ closest to nu and nv for a large integer n, and for a suitably small value $\epsilon > 0$ to let

$$t_H = n\mu(u) + a_u(n) g_u \qquad \text{and} \qquad t_V = n\mu(v) + a_v(n) g_v.$$

The integer n is chosen so that R^d is negligible compared to the minimum of n, $a_u(n)$, and $a_v(n)$ — this is possible because of the hypothesis that the sequences $a_u(n), a_v(n)$ go to ∞ — and so that the approximation (3) to the distributions of $T(nu), T(nv)$ by G_u, G_v (in particular, to their $(1-\epsilon)$th quantiles and their conditional expectations) is sufficiently accurate. For such n, if $\delta > 0$ is sufficiently small then for all edges e,

$$P\{X_e = 1\} > 1 - 2\delta - 2\epsilon, \tag{9}$$

and

$$E(Y_e - n\mu(u) \,|\, X_e = 1) < -4CR^d t_S \tag{10}$$

where C is the volume of the d-dimensional unit ball (thus, CR^d is, approximately, the number of edges of the lattice that intersect the ball B_v). Observe that (9) follows from the choice of t_H, t_V, as g_u, g_v are $(1-\epsilon)$th quantiles of G_u, G_v, and (10) from the fact that $\nu_u, \nu_v < 0$. Finally, choose t_S so large that $P\{X_v = 1\} > 1 - \delta$ for each vertex v (or, if the passage time distribution F has finite support, so that $P\{X_v = 1\} = 1$). Note that if $X_v = 1$ then the travel time between any two vertices in B_v is bounded by $2CR^d t_S$. In particular, if $X_v = X_{v'} = 1$ for both endpoints v, v' of an edge e, then the travel time from $\varphi(v)$ to $\varphi(v')$ is at most $4CR^d t_S + Y_e$.
The proof of Theorem 1 in the special case where the edge passage time distribution F has bounded support may now be completed. Assume that $\epsilon, \delta > 0$ are small enough that $1-2\epsilon-2\delta$

is no smaller than the critical value p_c for standard oriented bond percolation on the two-dimensional integer lattice. Then by (9) and Proposition 3, the percolation event Ω has positive probability. On this event, there is an infinite open oriented path γ starting at the origin that crosses the diagonal infinitely often. Consider now the image in $\mathbb{Z}^d$ of this path by the embedding φ: For each edge $e \in \gamma$, the traversal time $\tau(\gamma)$ is no larger than $Y_e + 4CR^d t_S$. Moreover, conditional on the realization $\{X_e\}$, the random variables Y_e, for $e \in \gamma$, are independent, with conditional distribution depending only on the orientation of e (vertices or horizontal), and with conditional expectations satisfying (10) above. Thus, by the SLLN, for (m, m) on the path γ the traversal time T_m of the segment of $\varphi(\gamma)$ from the origin to $\varphi(m, m)$ will satisfy

$$\liminf_{m \to \infty} T_m/m - n\mu(u) - n\mu(v) < 0 \tag{11}$$

Moreover, the liminf, when divided by n, remains bounded away from 0 as $n \to \infty$. As $n \to \infty$, the limiting directions of $x \approx nu$ and $y \approx nv$ approach u and v, respectively. Consequently, by the homogeneity and continuity of the inverse speed function $\mu(\cdot)$, the relation (4) follows for $\alpha = 1/2$.

Note. To prove the relation (4) in general, it suffices to prove it for $\alpha = 1/2$. This follows from the homogeneity of μ and the fact that any convex combination $\alpha u + (1-\alpha)v$ may be realized as $(u' + v')/2$ for suitable positive scalar multiples $u' = cu$ and $v' = c'v'$.

As has already been noted, if the edge passage time distribution F does not have bounded support then a modification of the argument is necessary, because in this case the Bernoulli random variables X_e and X_v need not be independent, and so Proposition 3 is not directly applicable. In this case, the random variables Y_e should be replaced by Y'_e, where Y'_e is defined as follows: Reset the edge passage time to t_S for all edges e that intersect one of the balls B_v, and let all other edge passage times be chosen from the distribution F. Now let Y'_e be the minimum traversal time $\tau'(\gamma)$ among all paths from z to z' that lie entirely in D_e, as before, and define

$$X'_e = 1 \quad \text{if} \quad Y'_e \leq t_H \quad \text{and } e \text{ is horizontal;} \tag{12}$$

$$X'_e = 1 \quad \text{if} \quad Y'_e \leq t_V \quad \text{and } e \text{ is vertical;} \tag{13}$$

and $X'_e = 0$ otherwise. The random variables X'_e, X_v are now mutually independent, and so Proposition 3 applies. Moreover, the random variables Y'_e are, conditional on the realization X'_e, X_v of the modified oriented percolation process, independent, with distribution depending only on the orientation of the edge e, and conditional means satisfying (10), at least when n is sufficiently large (since the modification of the edge passage times in the balls B_v will be swamped by the renormalization factors $a_u(n), a_v(n)$ for large n). The argument may now be completed in the same manner as in the special case discussed above.

References

[1] Benjamini, I., Kalai, G., and Schramm, O. (2003) First passage percolation has sublinear distance variance. To appear in *Annals of Probability.*

[2] Durrett, R., and Liggett, T. (1981) The shape of the limit set in Richardson's growth model. *Annals of Probability* **9** 186 – 193.

[3] Durrett, R. (1984) Oriented percolation in two dimensions. *Annals of Probability* **12** 999 – 1040.

[4] Grimmett, G. (1989) *Percolation.* Springer-Verlag, New York.

[5] Hammersley, J. M. and Welsh, D. J. A. (1965) First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. *Proc. Internat. Res. Sem., Statist. Lab., Univ. California, Berkeley, Calif.* 61 – 110.

[6] Johansson, K. (2000) Transversal fluctuations for increasing subsequences on the plane. *Probability Theory and Related Fields* **116** 445 – 456.

[7] Kesten (1986) Aspects of first passage percolation. *École d'été de probabilitéde Saint-Flour, XIV—1984* 125 – 264.

[8] Kordzakhia, G. (2003) The problem of coexistence in multi-type competition models. University of Chicago Ph. D. dissertation.

[9] Licea, C.; Newman, C. M.; Piza, M. S. T. (1996) Superdiffusivity in first-passage percolation. *Probability Theory and Related Fields* **106** 559–591.

[10] Licea, C.; Newman, C. M. (1996) Geodesics in two-dimensional first-passage percolation. *Annals of Probability* **24** 399 – 410.

[11] Liggett, T. (1985) *Interacting Particle Systems.* Springer-Verlag, New York

[12] Newman, C. M.; Piza, M. S. T. (1995) Divergence of shape fluctuations in two dimensions. *Annals of Probability* **23** 977 – 1005.

[13] Newman, C. M. (1995) A surface view of first-passage percolation. *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zurich, 1994)* 1017 – 1023.

[14] Pemantle, R.; Peres, Y. (1994) Planar first-passage percolation times are not tight. *Probability and phase transition (Cambridge, 1993)* 261 – 264.

[15] Russo, L. (1981) On the critical percolation probabilities. *Z. Warscheinlichkeitstheorie* **56** 229–237.