A simple proof of exponential decay in the two dimensional percolation model

Yu Zhang

December 25, 2008

Abstract

In 1980, Kesten showed the exponential decay of percolation probability in the subcritical phase for the two-dimensional percolation model. This result implies his celebrated computation that \( p_c = 0.5 \) for bond percolation in the square lattice, and site percolation in the triangular lattice, respectively. In this paper, we present a simpler proof for Kesten’s theorem.

1 Introduction and statement of results.

We may deal with the percolation model on the two-dimensional periodic lattice (see the detailed definition in Kesten (1982)). For simplicity, we select the triangular lattice without loss of generality, since we do not need to deal with the dual lattice separately. Consider site percolation on the triangular lattice. Each vertex of the lattice is open with probability \( p \) and closed with probability \( 1 - p \), and the sites are open independently of each other. We will realize the triangular lattice with vertex set \( \mathbb{Z}^2 \). For a given \( (x, y) \in \mathbb{Z}^2 \), its nearest neighbors are defined as \( (x \pm 1, y) \), \( (x, y \pm 1) \), \( (x + 1, y - 1) \), and \( (x - 1, y + 1) \). Bonds between neighboring or adjacent sites therefore correspond to vertical or horizontal displacements of one unit, or diagonal displacements between the two nearest vertices along a line making an angle of 135° with the positive x-axis. Recall that the triangular lattice may also be viewed with sites as hexagons in a regular hexagonal tiling of the plane. The corresponding probability measure on the configurations of open and closed sites is denoted by \( P_p \). We also denote by \( E_p \) the expectation with respect to \( P_p \).

A path from \( u \) to \( v \) is a sequence \( (v_0, e_1, v_1, \ldots, v_i, e_{i+1}, v_{i+1}, \ldots, v_l) \) with distinct vertices \( v_i \) \((0 \leq i \leq n)\) and \( v_0 = u \) and \( v_n = v \) and with bonds \( e_i \) between \( v_i \) and \( v_{i+1} \). If \( u = v \), the path

AMS classification: 60K 35.

Key words and phrases: percolation and critical probability.

\*Research supported by NSF grant DMS-0405150.
is called a *circuit*. If all of the sites in a path are open, the path is called an *open path*. Given a rectangle $[-n, n] \times [-m, m]$, a *left-right open crossing* is a path $(v_0, e_1, v_1, \ldots, v_i, e_{i+1}, v_{i+1}, \ldots, v_t)$ in $[-n, n] \times [-m, m]$ such that all of its vertices inside $(-n, n) \times (-m, m)$ are open except $v_0$ and $v_t$, which are at the left and at the right of the rectangle, respectively. Similarly, we can define a *top-bottom open crossing*. We denote by $LR([−n, n] \times [−m, m])$ and $TB([−n, n] \times [−m, m])$ the events that there exist the left-right and the top-bottom open crossings, respectively. We may replace the open paths with closed paths to have the events of $LR^*([-n, n] \times [-m, m])$ and $TB^*([-n, n] \times [-m, m])$, respectively. We call them the *left-right closed crossing* and the *top-bottom closed crossing*, respectively.

There might be many open crossings. For each open crossing $\Gamma$, it divides $[-n, n] \times [-m, m]$ into two parts: the top part $T(\Gamma)$ and the bottom part $B(\Gamma)$, including the top and the bottom of $[-n, n] \times [-m, m]$, respectively. We also denote by $T^o(\Gamma)$ and $B^o(\Gamma)$ the interiors of $T(\Gamma)$ and $B(\Gamma)$, respectively. If there is more than one left-right open crossing, we select an open crossing with the smallest vertex set $B(\Gamma)$. We call the open crossing the lowest crossing. Without loss of generality, we still denote by $\Gamma$ the lowest open crossing. By the definition of the lowest crossing, it can be obtained (see page 317 in Grimmett (1999) or Proposition 2.3 in Kesten (1982)) that the event of $\{\Gamma = \gamma\}$ for some fixed left-right crossing $\gamma$ only depends on the open or closed vertices in $B(\gamma)$. This property is said to be the independent property of the lowest crossing.

If $\Gamma$ is the lowest open crossing, then for each $v \in \Gamma$, it is well known (see Proposition 2.2 in Kesten (1982)) that there exists a closed path from $v$ (not including $v$) to the bottom of $[-n, n] \times [-m, m]$. By this observation, for each $v \in \Gamma$, there exist two disjoint open paths $\Gamma_1$ and $\Gamma_3$ with

$$\Gamma_1 \cup \{v\} \cup \Gamma_3 = \Gamma,$$

from $v$ to the left and to the right of $[-n, n] \times [-m, m]$, respectively. In addition, there exists a closed path $\Gamma_2$ in $B^o(\Gamma)$ from $v$ to the bottom of $[-n, n] \times [-m, m]$. On the other hand, by using Proposition 2.2 in Kesten (1982) again, if there exist the three paths $\Gamma_i$ for $i = 1, 2, 3$ at $v$, and $v$ is open, then $v$ is on the lowest crossing. This is called the three-arm-path argument for each vertex on the lowest crossing $\Gamma$. Similarly, the three-arm-path argument holds for the left-most top-bottom crossing.

We may generalize these arguments into a circuit enclosed by a path. Let $\Delta$ be an open set surrounded by a circuit $\partial \Delta$. We select four vertices $v_i$ for $i = 1, 2, 3, 4$ from the circuit. Let $L$ (left), $T$ (top), $R$ (right), and $B$ (bottom) be the paths along $\Delta$ clockwise from $v_1$ to $v_2$, from $v_2$ to $v_3$, from $v_3$ to $v_4$, and from $v_4$ to $v_1$, respectively. With these paths, we can define the events $LR(\Delta)$ and $TB^*(\Delta)$ in the same way that we did for a rectangle. In addition, we can also define the lowest left-right, and the left-most top-bottom open or closed crossings. We denote by $\Gamma_{LR}(\Delta)$ and $\Gamma_{TB}^*(\Delta)$ the lowest left-right open, and the left-most top-bottom closed crossings. By the same discussion above, the independent property of the lowest crossing and the three-arm-path argument still hold for $\Gamma_{LR}(\Delta)$ and $\Gamma_{TB}^*(\Delta)$. In fact, Kesten (1982) discussed these topology properties precisely in the circuit as we defined.
For each $v \in [-n, n]^2$, we say there are three arm paths from $v$, as we mentioned, if there are two disjoint open paths $\Gamma_1$ and $\Gamma_3$ in $[-n, n]^2$ from $v$ to the left and to the right of $[-n, n]^2$, and there exists a closed path $\Gamma_2$ from $v$ to the bottom of $[-n, n]^2$, respectively. Also, $v$ is open. Moreover, if there exists an additional closed path $\Gamma_4$ in $T^o(\Gamma)$ from $v$ to the top boundary of $[-n, n]^2$, we say there are four arm paths at $v$. If there are four arm paths at $v$, $v$ is said to be a pivotal vertex of open crossing $LR([-n, n]^2)$. Let $N_n$ be all the pivotal vertices in $[-n, n]^2$.

The open cluster of the vertex $x$, $C(x)$, consists of all vertices, that including $x$, that are connected by open paths. For any collection $A$ of vertices, $|A|$ denotes the cardinality of $A$. We choose $0$ as the origin. The percolation probability and the mean size of the open cluster are denoted by

$$\theta(p) = P_p(|C(0)| = \infty) \quad \text{and} \quad \chi(p) = E_p(|C(0)|),$$

and the critical probabilities are defined by

$$p_c = \sup \{ p : \theta(p) = 0 \} \quad \text{and} \quad p_T = \sup \{ p : \chi(p) < \infty \}.$$ 

Similarly, we denote by $C^*(x)$ the closed cluster including $x$. With these definitions, the crucial step in Kesten’s paper (1980) is to estimate $E_p(N_n, LR([-n, n]^2))$ in the following theorem. In this paper, we will present a simpler proof for his estimate.

**Theorem 1.** If $p \leq 0.5$, then there exists $\alpha > 0$ such that for all $n$,

$$n^\alpha \leq E_p \left( N_n \mid LR\left( [-n, n]^2 \right) \right). \quad (1.1)$$

With Theorem 1, we will have the following corollary.

**Corollary 2.** If $p < 0.5$, then there exist constants $C_i = C_i(p)$ for $i = 1, 2$ such that

$$P_p(C(0) \cap \partial [-n, n]^2 \neq \emptyset) \leq C_1 \exp(-C_2n) \quad (1.2)$$

and

$$P_{1-p}(C^*(0) \cap \partial [-n, n]^2 \neq \emptyset) \leq C_1 \exp(-C_2n), \quad (1.3)$$

where $\partial [-n, n]^2$ is the boundary vertex set of $[-n, n]^2$.

**Remark.** If (1.2) holds, it follows from Theorem 5.1 in Kesten (1982) that

$$P_p(|C(0)| \geq n) \leq C_1 \exp(-C_2n).$$

For more than two decades since 1959, one of the most important discoveries in the history of the percolation model was the rigorous determination of $p_c = 0.5$ for the square
lattice and the triangular lattice. Harries (1960) proved that $p_c \geq 0.5$. The precise lower bound of $p_c$ seems much harder to achieve. After 20 years, by the estimate in Theorem 1, Kesten (1980) finally showed that $p_c = 0.5$. In this paper, we present a proof by using Theorem 1 to show $p_c = 0.5$.

**Corollary 3.** $p_c = p_T = 0.5$.

**Remark.** The same argument can be carried out to show Theorem 1 if $p \leq p_c$, and Corollary 2 if $p < p_c$ for the percolation model in the two-dimensional periodic lattice. In addition, the same argument can also be carried out to show that $p_c = p_T = 0.5$ for the bond percolation model in the square lattice.

## 2 Proofs of theorems and corollaries.

Before the proofs of the Theorems, we introduce a lemma by Russo (1978) and Seymour and Welsh (1978).

**RSW lemma.** If $P_p(LR([0, n]^2)) \geq \delta > 0$, then for each integer $k$, there exists a positive constant $C_3 = C_3(k, \delta)$ such that

$$P_p(LR([0, kn] \times [0, n])) \geq C_3. \quad (2.1)$$

Note that by symmetry, we know that

$$P_{0.5}(LR([0, n]^2)) = 0.5 \text{ and } P_{0.5}(LR^*([0, n]^2)) = 0.5$$

for each $n$, so by the RSW lemma,

$$P_{0.5}(LR([0, kn] \times [0, n])) \geq C_3 \text{ and } P_{0.5}(LR^*([0, kn] \times [0, n])) \geq C_3. \quad (2.2)$$

With (2.2) and the FKG inequality, we can directly show that there exists a closed circuit in an annulus with a positive probability. Thus the following lemma can be directly obtained by this probability estimate (see Theorem 11.89 in Grimmett (1999)).

**Lemma 1.** There exists $C_4 > 0$ such that

$$P_{0.5}(C(0) \cap \partial[-n, n]^2 \neq \emptyset) \leq n^{-C_4}.$$

Now we show Theorem 1 by using Lemma 1.
Proof of Theorem 1. We first estimate the pivotal sites when $p = 0.5$. Let (see Fig 1.) $L_n(\gamma)$ be the event that the lowest open crossing on $[-n, n]^2$ is $\gamma$ for a fixed crossing $\gamma$. On $L_n(\gamma)$, the lowest crossing $\gamma$ has to stay in $[-n, n]^2$. Next, for each fixed lowest crossing on $[-n, n]^2$, let (see Fig. 1) $D_1(\gamma)$ be the event that there exists two open paths: one from the top of $[-n/2, 0] \times [-n, 3n/2]$ to $\gamma$ inside $T^\circ(\gamma) \cap [-n/2, 0] \times [-n, 3n/2]$, and the other one from the left to the right in $[-n, 0] \times [n, 3n/2]$. In addition, let $D_2(\gamma)$ be the event that there exists a closed path from the top boundary of $[0, n/2] \times [-n, 3n/2]$ to $\gamma$ inside $T^\circ(\Gamma) \cap [0, n/2] \times [-n, 3n/2]$. By independent property, for each fixed crossing,

$$P_p(D_1(\gamma) \cap D_2(\gamma) \mid L_n(\gamma)) = P_p(D_1(\gamma) \cap D_2(\gamma)).$$

If there exists an open path from the top to the bottom of $[-n/2, 0] \times [-n, 3n/2]$ inside of $[-n/2, 0] \times [-n, 3n/2]$, then there exists an open path from the top of $[-n/2, 0] \times [-n, 3n/2]$ to $\gamma$ inside $T^\circ(\gamma) \cap [-n/2, 0] \times [-n, 3n/2]$. With this observation, the FKG inequality, and (2.2), there exists a constant $C_4 > 0$ such that

$$P_{0.5}(D_1(\gamma)) \geq C_4.$$

The same argument implies that

$$P_{0.5}(D_2(\gamma)) \geq C_5$$

for some constant $C_5 > 0$. Note that for each fixed $\gamma$, $D_1(\gamma)$ and $D_2(\gamma)$ are independent, so there exists $C_6 > 0$ such that

$$P_{0.5}(D_1(\gamma) \cap D_2(\gamma) \mid L_n(\gamma)) \geq C_6. \tag{2.3}$$
On \( L_n(\gamma) \), the lowest crossing on \([-n, n]^2\) is \( \gamma \). Note that the boundary of \( T(\gamma) \) is a circuit enclosed by the following four pieces (see Fig. 1): the bottom \((B = \gamma)\), the top \(T\) (the top boundary of \([-n, n]^2\)), the left \(L\) (the part of the left boundary of \([-n, n]^2\)), and the right \(R\) (the part of the right boundary of \([-n, n]^2\)).

With these \(L, T, R\), and \(B\), we consider the left-most top-bottom closed crossing in \(T(\gamma)\). On \(L_n(\gamma)\), if \(\mathcal{D}_1(\gamma) \cap \mathcal{D}_2(\gamma)\) occurs, this left-most top-bottom closed crossing \(\Gamma_4\) (see Fig. 1) exists in \(T^c(\gamma)\) with a starting vertex \(v_0 \in [-n/2, n/2] \times [-n, n]\) (not included on \(\Gamma_4\)) at \(\gamma\). On \(L_n(\gamma)\), if \(\mathcal{D}_1(\gamma) \cap \mathcal{D}_2(\gamma)\) occurs, we denote by \(\mathcal{E}_{v_0}(\gamma)\) the event that there exists the lowest left-right open crossing \(\gamma\) passing through \(v_0\), and there exists the left-most top-bottom closed crossing \(\Gamma_4\) from \(\gamma\) to the top boundary of \([-n, n] \times [-n, 3n/2]\) with the starting vertex \(v_0 \in [-n/2, n/2] \times [-n, n]\). Since \(\Gamma_4\) is the left-most closed crossing, by the three-arm-argument there exists an additional open path \(\Gamma_5\) inside \(T^c(\gamma)\) from a neighbor of \(v_0\), denoted by \(v_1\), to the left of \([-n, n] \times [-n, 0, 3n/2]\) (see Fig. 1). Here \(\Gamma_5\) includes \(v_1\).

In summary, if \(\mathcal{E}_{v_0}(\gamma)\) occurs, there are four disjoint paths from \(v_0\) (not including \(v_0\)): two open paths \(\Gamma_1\) and \(\Gamma_3\) from \(v_0\) to the left and to the right of \([-n, n] \times [-n, 3n/2]\), and two closed paths \(\Gamma_2\) and \(\Gamma_4\) from \(v_0\) to the top and to the bottom of \([-n, n] \times [-n, 3n/2]\), respectively. Also, \(v_0\) is open. In addition to these four arm paths, there exists an open path \(\Gamma_5\) in \(T^c(\gamma)\) from \(v_1\) to the left of \([-n, n] \times [-n, 3n/2]\) and \(v_1\) is open. Note that \(v_1\) is a neighbor of \(v_0\). Note also that if \(v_0\) is fixed, then there are at most nine choices for choosing \(v_1\). Let \(N([-n, n] \times [-n, 3n/2])\) be the number of pivotal sites for the crossing \(LR([-n, n] \times [-n, 3n/2])\). Therefore, for fixed crossing \(\gamma\), by (2.3), Reimer’s inequality (2000), translation invariance, Lemma I, and independent property,

\[
C_6 \leq P_{0.5} (\mathcal{D}_1(\gamma) \cap \mathcal{D}_2(\gamma) | L_n(\gamma)) \\
= P_{0.5} \left( \bigcup_{v_0 \in [-n/2, n/2] \times [-n, n]} \mathcal{E}_{v_0}(\gamma) \mid L_n(\gamma) \right) \\
\leq \sum_{v_0 \in [-n/2, n/2] \times [-n, n]} P_{0.5} (\mathcal{E}_{v_0}(\gamma) \mid L_n(\gamma)) \\
\leq 9 \sum_{v_0 \in [-n/2, n/2] \times [-n, n]} P_{0.5} (\mathcal{C}(0) \cap \partial [-n/2, n/2]^2 \neq \emptyset) \\
\cdot P_{0.5} (\exists \text{ four arm paths at } v_0 \text{ for } [-n, n] \times [-n, 3n/2] | L_n(\gamma)) \\
\leq 9n^{-C_4} E_{0.5} \left( N([-n, n] \times [-n, 3n/2, 3n] | L_n(\gamma)) \right). 
\tag{2.4}
\]

On \(L_n(\gamma)\), the lowest open crossing on \([-n, n] \times [-n, 3n/2]\) stays inside \([-n, n]^2\). Thus, on \(L_n(\gamma)\), each pivotal site for the left-right open crossing of \([-n, n] \times [-n, 3n/2]\) is also a pivotal site for the left-right open crossing of \([-n, n]^2\). In other words, for each \(p\),

\[
E_p \left( N([-n, n] \times [-n, 3n/2] | L_n(\gamma)) \right) \leq E_p \left( N_n | L_n(\gamma) \right). 
\tag{2.5}
\]
Together with (2.4) and (2.5), we have for each fixed crossing $\gamma$,
\[
 n^\alpha \leq \mathbb{E}_{0.5} (N_n \mid L_n(\gamma)). \tag{2.6}
\]

Now we show Theorem 1 by (2.6).
\[
 E_p \left( N_n; LR([-n, n]^2] \right) = \sum_\gamma E_p \left( N_n \mid L_n(\gamma) \right) P_p(L_n(\gamma)). \tag{2.7}
\]

On $L_n(\gamma)$, if there exist four arm paths at $v$, then by the three-arm-path argument, $v$ is on the lowest crossing. Therefore, $N_n$ is the number of vertices $\{v\} \subset \gamma$ such that there exist closed paths inside $T^0(\gamma)$ from $v$ (not including $v$) to the top of $[-n, n]^2$. For each fixed crossing $\gamma$, let $V_n(\gamma)$ be the vertices of $\{v\}$ above. By the independence property of the lowest crossing, these closed paths only depend on the configurations on $T^0(\gamma)$:
\[
 E_p \left( N_n \mid L_n(\gamma) \right) = E_p \left( V_n(\gamma) \right).
\]

Note that $E_pV_n(\gamma)$ is decreasing in $p$ for each fixed crossing $\gamma$. Therefore,
\[
 E_{0.5} \left( N_n \mid L_n(\gamma) \right) = E_{0.5} \left( V_n(\gamma) \right) \leq E_p \left( V_n(\gamma) \right) = E_p \left( N_n \mid L_n(\gamma) \right). \tag{2.8}
\]

By (2.6), (2.7), and (2.8), for all $p \leq 0.5$,
\[
 n^\alpha P_p(LR([-n, n]^2]) \leq E_p \left( N_n; LR([-n, n]^2] \right). \tag{2.9}
\]

Theorem 1 follows. \Box

**Remark.** Kesten, Sidoravicius, and Zhang (1998) gave a precise order of the probability estimate for the five arm paths. The proof is quite long.

If we denote that $N^*_n$ be the pivotal sites for a closed crossing in $[-n, n]^2$, then by symmetry and Theorem 1, we have the following Corollary.

**Corollary 4.** If $q \geq 0.5$, then
\[
 E_q \left( N^*_n; L^*([n, n]^2] \right) \geq n^\alpha. \tag{2.10}
\]

**Proof of Corollary 2.** By Theorem 1 and Russo’s formula (see (2.30) in Grimmett (1999)), note that $LR([-n, n]^2]$ is an increasing event, so there exist $C_i = C_i(p)$ for $i = 7, 8$ such that for $p < 0.5$,
\[
 P_p(LR([-n, n]^2]) \leq \exp \left(- \int_p^{0.5} E_p(N_n \mid LR([-n, n]^2]) \right) \leq C_7 \exp(-C_8 n^\alpha). \tag{2.11}
\]
By (2.10) and symmetry, if \( q > 0.5 \), then
\[
P_q(LR^*([−n, n]^2)) \leq C_7 \exp(−C_8 n^α). \tag{2.12}
\]
Note that if \(|C(0)| \geq n\), then there exists an open path from the origin to \(\partial[−\sqrt{n}, \sqrt{n}]^2\). By (2.12), symmetry and the FKG inequality, there exist \(C_i = C_i(p)\) for \(i = 9, 10\) such that for \(p < 0.5\),
\[
P_p(|C(0)| \geq n) \leq [P_p(LR([-\sqrt{n}, \sqrt{n}]^2))]^{1/2} \leq C_9 \exp(−C_{10} n^α/2). \tag{2.13}
\]
By (2.13),
\[
p_T \geq 0.5. \tag{2.14}
\]
Corollary 2 also follows from (2.13) and a simple computation (see Theorem 5.4 in Grimmett (1999)). □

**Proof of Corollary 3.** By Lemma 1, for each \(n\),
\[
P_{0.5}(|C(0)| = \infty) \leq P_{0.5} \left(C(0) \cap \partial[−n, n]^2 \neq \emptyset\right) \leq n^{-C_4}. \tag{2.15}
\]
Thus, \(θ(0.5) = 0\), so
\[
p_c \geq 0.5. \tag{2.16}
\]
Now we assume that \(p_c > 0.5\) and select \(0.5 < q < p_c\). With this assumption, by (2.2),
\[
P_q(LR^*([−n, n]^2)) \geq C_3 > 0. \tag{2.17}
\]
Since (2.12) and (2.17) cannot hold together for large \(n\), the contradiction tells us that \(p_c \leq 0.5\). Together with (2.16), we have \(p_c = 0.5\). Note that \(p_T \leq p_c\), so by (2.14), \(p_T = 0.5\). Therefore, Corollary 3 follows. □

**References**

Grimmett, G. (1999). *Percolation*. Springer-Verlag, New York.
Harries, E. (1960). A lower bound for the critical probability in a certain percolation process. *Proceedings of the Cambridge Philosophical Society* 56 13–20.
Kesten, H. (1980). The critical probability of bond percolation on the square lattice equals 1/2. *Comm. Math. Phys.* 74 41–59.
Kesten, H. (1982). *Percolation Theory for Mathematicians*, Birkhauser, Boston.
Kesten, H., Sidoravicius, V. and Zhang, Y. (1998). Almost all words are seen in critical site percolation on the triangular lattice. *Electron. J. Probab.* 3 1–75.
Reimer, D. (2000). Proof of the van den Berg-Kesten inequality. *Combin. Probab. Compute* 9 27–32.
Russo, L. (1978). A note on percolation. *Z. Wahrsch. verw. Gebiete* 43 39–48.
Seymour, P. D. and Welsh, D. J. A. (1978). Percolation probabilities on the square lattice. In *Advances in Graph Theory* (B. Bollobas ed.) 227–245. Ann. of Discrete Math. 3, North-Holland, Amsterdam.

Yu Zhang  
Department of mathematics  
University of Colorado  
Colorado Springs, CO 80933  
yzhang3@uccs.edu