September 1982 Report No. STAN-CS-82-933

An Algorithmic Method for Studying Percolation Clusters

by

Shmuel T. Klein and Eli Shamir

Department of Computer Science

Stanford University
Stanford, CA 94305

September 1982

An Algorithmic Method for Studying Percolation Clusters

by
Shmuel T. Klein
The Weizmann Institute
and
Bar-Ilan University

and

Eli Shamir
Hebrew University
and
Stanford University

Research of the 2nd author
partially sponsored by the
Systems Development Fund contract

1. Introduction

Percolation Theory originated in problems of fluid flow in random media, e.g. molecules penetrating a porous solid or a disease infecting a population. The applications in Statistical Mechanics brought the main developments in percolation theory. How does a drastic phase-transition of a system occur across a narrow band of the system parameters (e.g. temperature, density)? The mathematical explanation goes as follows: A physical system is a probability space of microscopic configuration (say N particles occupying sites in a certain lattice). The space is parametrized by a certain controllable quantity p (or $p_1, p_2, \ldots$). The physically observable phase is determined by some intrinsic property Q of the configuration. One show that

Prob {Q} = Prob {Configuration has property Q},

which is a function of p, undergoes a drastic increase from ϵ to $1 - \epsilon$ across a narrow band of p-values. Usually ϵ depends on N, the number of particles, and c(N) = o(N).

It was noticed that the situation here is the precise analog of random-graph properties having a sharp threshold, where the parameter p measures edge-density. Moreover, the phase-determining properties are usually size, number and shape of particle-clusters. These are essentially like connectivity components in graphs, for which sharp thresholds apply. Here we argue that methods from one area will be useful for the other.

Our main contribution is that a certain configuration-dependent quantity, the boundary-body density ratio of "giant" clusters, is sharply determined by the parameter p (a full-site frequency) of the system. Therefore, as physicists also argue from microscopic considerations, it has a physical significance (related to energy). The way we prove it is useful in studying random-graph algorithms [Sh 82]. A random procedure ALG gives an intertwined construction of the configurations and giant clusters in them. Following ALG we are able to compute the required ratio.

We believe this approach, and percolation theory in general, is useful in the study of large technologically designed systems, e.g. networks of communicating processors with a suitable geometry.

2. Percolation Models and Clusters

We consider an interaction-f&e model. The underlying medium is an infinite lattice L of some finite dimension d. The lattice L consists of a countable set of vertices, called *sites*, linked by edges, called *bonds*.

In the *site-percolation* model, all the bonds represent fixed (open) connections, while each site is assigned, independently of the others, values "F" or "E" with probabilities

p for "F" (full), $1 - $ p for "E" (empty).

This results in a probability space $\Omega = \Omega(L, p)$, its elements are called *configurations*.

In the bond **perculation** model, the sites represent fixed (open) vertices while the bonds, like edges in a random graph, occur independently with probability p. By the line-graph construction, i.e. passing to a lattice *L'* whose sites are the (midpoints of the) bonds of L, we see that it suffices to consider the site-percolation model, which is more general. For a fluid model, "F" represents open, "E" represents closed to the flow. This leads naturally to the next definition.

A *connection* between "F"-**marked** sites a and b is a path of (alternating) bonds and "F"-marked sites, leading from a to b (so along a path the fluid can flow). A *cluster C* is **a** maximally connected set of "F"-sites, the size $|C|$ is the number of sites in C.

Denote by R,(p) the probability that a particular site belongs to a cluster of size at least n. Clearly $R_n(p) \geq R_{n+1}(p)$. *So*

$$R(p) = \lim_{n\to\infty} R_n(p) \quad \text{exists for each p.}$$

This *R(p)* is the probability of any particular site b to belong to an *infinite* ("giant") *cluster* (the fluid in b can flow to an unbounded distance).

It can bc shown that *R(p)* is a non-decreasing function of p. Clearly *R(0)* = 0, R(1) = **1**. Let

$$p_c = p_c(L) = glb(p|R(p) > 0)$$

p_c is called the critical percolation value. Again it is easy to show that for $p > p_c$ an infinite cluster exists in ω, with probability 1 in $\Omega(L,p)$.

The numerical values of *R(p)* and p_c depend on the dimension and the geometry of the' lattice. Deriving them from theoretical calculation is usually hard. One can estimate them by experiments (physical simulations) or by Monte Carlo methods (computer simulations).

3. The Number of Infinite Clusters

Let N,(w) denote the number of infinite clusters in the configuration ω. It was shown [NS 81] that N_∞ has a constant value with probability **1.** This value can be 0 if (p < p_c) *or* **1** or ∞. The *uniqueness conjecture* is:

$$p > p_c \Rightarrow N_\infty = 1 \text{ with probability 1.}$$

The general belief is that this must be true, some even speak about "the" infinite cluster without noticing that "the" is **a** problem. In fact, uniqueness was proved only for planar lattices, and is wide open in higher dimensions.

Naturally, the aspect of shape of infinite clusters is related to that of number and size. To discuss it, we need a structure of a monotone increasing sequence of bounded domains

$$Q(0) \subseteq Q(1) \subseteq \ldots \subseteq Q(n) \subseteq \ldots, \bigcup_{i=1}^{\infty} Q(i) = L$$

which exhaust the lattice L. Usually some regularity is required. The typical sequence for the cubic grid (in any dimension) is Q(n) = the cube of edge length 272 + **1** centered **at** $(\frac{1}{2}, \ldots, \frac{1}{2})$.

The density of a set of sites C **is**

$$\text{dens}(C) = \liminf_{n \to \infty} \frac{|C \bigcap Q(n)|}{|Q(n)|}$$

Clusters with dens(C) $>$ 0 are called called *dense,* those with density 0 are filamentary. It was shown in [NS **82**] that, with probability **1** there is at most one dense cluster. We can further show [KS **81**]

existence of a dense cluster implies uniqueness,

i.e. $N_\infty = \infty \Rightarrow$ all clusters are filamentary.

Our main result here will show that all infinite clusters have a similar "shape." Intuitively this adds credibility to the uniqueness conjecture because such regimented behavior of **a** configuration seems to have vanishing probability.

4. The Boundary-Body (BB) density quotient

The boundary of a cluster C, denoted by ∂C, is the set of sites such that $x \neg\epsilon C$ **but** x is adjacent to some site $y \epsilon C$. Let

$$s_n = |C \bigcap Q(n)|, \quad t_n = |\partial C \bigcap Q(n)|$$

We want to study t_n/s_n. An asymptotic relation

$$t_n = \frac{1-p}{p} s(n) + 0(s_n^\varsigma), \varsigma < 1, n \to 00$$

will be proved for infinite clusters where ς is essentially $\frac{1}{2}$. This means that C is a highly ramified set, unlike an expanding ball or cube, its surface grows linearly with its volume.

A lively history of this BB density quotient is related in [St **79**]. In [ADS **80**] it **is** proved that $\frac{<t_s>}{s} \to \frac{1-p}{p}$ as $s \to \infty$, where $< t_s >$ is the *expected* boundary size for finite clusters of size s. [H 79] gave a rather obscure argument why the asymptotic limit for clusters of finite size should be equal to the BB density quotient of "the" infinite cluster.

Our proof, along the lines described below, was found in 1980, following a lecture of Professor C. Domb in Jerusalem. It was commuicated to Newman and Schulman, who then gave another proof [NS] *under* the assumption *of uniqueness,* using ergodicity of the translations in regular lattices. The claims about the remainder term s^ς are quite confused in the literature [St **79**]. How were the simulations made? Is ς claimed to depend on the lattice, the dimension, or on $p - p_c$? Our result gives $s = \frac{1}{2}$ for the infinite cluster, which suggests that the same should hold for the finite cluster asymptotics.

THEOREM 1. *Let $p > p_c$. with probability one in $\Omega(L,p)$, the following relation holds for any infinite cluster C*

$$t_n = \frac{1-p}{p} s(n) + O(s_n)(\log \log s_n)^{\frac{1}{2}}, \ n \to \infty.$$

The computation of the boundary-body density quotient for an infinite cluster is based on a probabilistic procedure to construct ω, which we call ALG. Each $\omega \epsilon \Omega$ is determined by a choice of *marking* ("F" with probability p or "E" with probability $q = 1 - p$) for the random variable associated to each site. The order in which the marking is carried out is immaterial, and we can use it to our advantage; in particular we first "construct" an infinite cluster and its boundary, and complete then (in any order) the marking of those sites unmarked by ALG.

For simplicity, we treat a cubic lattice in R^d. Let Q(n) be the system of cubes exhausting L and Q'(n) = Q(n) \ $Q(n - 1)$ the *shell* of outermost sites of Q(n).

The procedure ALG consists of steps numbered 1,2,.. . . Up to step i, we mark sites only in $Q(i)$. Step i is concluded when we have marked all the sites in Q(i) which are adjacent to sites marked "F." Now *if* the external shell $Q'(i)$ has no site marked "F," *then* we know that the current cluster we follow is completely enveloped by its boundary, and -we pick an innermost unmarked site for marking, attempting to start a new cluster.

Else, there is a site x marked "F" in $Q'(i)$, we enter step $i+1$ by marking the neighbors of x in $Q'(i + 1)$ and continue, by a sequence of sub-steps, to mark all remaining sites in $Q(i + 1)$ until step $i + 1$ is concluded.

The MARKING procedure **ALG**

i - index of steps
j - index of sub-steps .
K - integer satisfying that all the finite clusters are included in $Q(K)$
h - index of cluster being enlarged
Mark (h,S) - is a routine, assigning to each of the elements of the set (of sites) S, the value h with probability p, or 0 with probabiity $q = 1 - p$

(A) [Initialization]
A.1 $i := 0$; $k := 0$; $h := 1$;
A.2 Mark (1, Q(0));

(B) [Next step]
B.1 *if* there, is a site marked h in $Q'(i)$
B.2 *then* [continue to enlarge cluster *h]*
B.3 $i := i + 1$; j := 0;
B.4 $goto(D)$;

(C) [No site was marked h in $Q'(i)$, hence cluster h is finite; search for next cluster]
C.1 $h := h + 1$;
Repeat
c.2 i' :=largest integer satisfying that all the sites in $Q(i')$ have been marked;
c.3 $i := i' + 1$;
c.4 x :=random unmarked site in Q(i);
C.5 Mark *(h, {x});*
C.6 *until* one site has been marked *h;*
C.7 $j := 0$; $K := max(K, i - 1)$;

(D) [next sub-step]
D.1 $j := j + 1$;
D.2 $R(i, j, h)$:= set of unmarked sites in $Q(i)$,
having at least one neighbor in cluster h;
D.3 *if* $|R(i, j, h)| > 0$
D.4 *then* Mark $(h, R(i, j, h))$
D.5 $goto(D)$
D.6 else *goto* (B);

5. Analysis of ALG

CLAIM: If there is an infinite cluster, ALG will lock on it after *a finite number of steps.*

PROOF: Let C_h denote the cluster, an element of which is chosen in the h-th passage through block (C), in C.4.

After the execution of block (C), we shall return to this block if and only if C_h happens to be finite, for *if* C_h is infinite, then

$$\exists n_o(n \geq n_o \Rightarrow C_h \bigcap Q'(n) \neq \emptyset$$
$$\wedge n < n_o \Rightarrow C_h \bigcap Q'(n) = 0) \qquad 0^*$$

$(Q(n_o - 1)$ is the largest cube centered in $(\frac{1}{2}, \ldots, \frac{1}{2})$ not intersecting $C_h)$; denote the value of i after the h-th execution of (C) by i_o, then

$$C_h \bigcap Q'(i_o) \neq \emptyset$$

hence $(*)$ implies $i_o \geq n_o$ and

$$\forall i \geq i_o \quad C_h \bigcap Q'(i) \neq \emptyset$$

. so that the condition in B.1 remains true in all the subsequent steps, thus (C) will not be executed again.

If C_h is finite, there is a minimal cube $Q(h_o)$ which contains the entire cluster, we get a negative answer in B.1 for $i = h_o + 1$ and ALG passes to (C).

Suppose that there is an infinite cluster A; once ALG starts marking sites of A, there will be no further passage through (C), and since only in this block *h* is increased (i.e. we choose a new cluster), ALG will not leave A any more. Thus it is enough to show that after a finite number of executed statements, some site in *A* has been marked.

Let n_o be the smallest integer for which $Q(n_o) \bigcap A \neq \emptyset$, then after at most $|Q(n_o)|$ passages through block (C) a site of *A* will be marked, as implied by the conditions in C.2 and C.6. Knowing that between two consecutive executions of (C) ALG enlarges a finite cluster C_h, the number of statements up to the first marking in *A* is bounded by

$$O(|Q(n_o)| \cdot \max_{h < |Q(n_o)|} (|C_h|)).$$

Thus ALG finds the infinite cluster, if there is one, and we obtain a correct generation of all the configurations in Ω by marking afterwards the sites which remained unmarked. If in some w constructed, no infinite cluster is obtained, then the whole lattice is already marked by ALG itself; for p > p_c, the set of such w has probability 0. This concludes the proof of the CLAIM. We proceed to the computation of the BB density quotient.

ALG scans the lattice in a certain order, which yields for a fixed configuration w, a sequence $U_1, U_2, \ldots$ where each U_i is a pair of random variables $(S(U_i), V(U_i))$. $S(U_i) \epsilon R^d$ is the i-th site marked. by ALG, $V(U_i)$ the value ("E" or "F") which $S(U_i)$ was assigned. By the definition of ALG, $\{V(U_i)\}_{i=1}^{\infty}$ is a Bernoulli sequence.

But in order to prove the theorem about the infinite cluster's boundary-to-body ratio, we need a rearrangement of our r.v.'s. Since this is a density quotient of two infinite sets, we have to exhaust the lattice with the sequence of domains Q(n), to assure that the quotient value is indeed a property. of the infinite cluster, and is not a consequence of a perhaps sophisticated way of counting the sites.

The easiest way of rearranging the marked sites is to enumerate them shell by shell and in each shell to proceed in a simple geometric order, e.g. passing over the shell by a sequence of adjacent sites. This will not do, since randomness is destroyed. For example, if Y_1 denotes the first element of the above sequence (Y_1 is one of the *{U..})*, then if we started scanning the first shell on an unmarked site, the first site we encounter is in ∂C and is certainly marked "E", so that

$$P(V(Y_1) = \text{"}E\text{"}) = (1-p)R + (1-R) \neq 1 - p$$

where R is the probability of a site being marked by ALG.

What happens'here is that the index of a variable of this rearrangement carries more information than a random sequence should, e.g. the event $\{S(Y_i) = S(U_k)\}$ depends on the whole sequence $\{U_i\}$ since it includes the event: "after *all* the sites have been marked, there are exactly $i - 1$ marked sites preceeding (in the given order) the one marked by U_k", whereas we are interested in having this event depend only on $U_1, \ldots, U_{k-1}$.

Therefore we choose a rearrangement which is a compromise between the geometric and random generation order: we form the blocks Y^k, $k = 1, 2, \ldots$, of marked sites belonging to the same shell *Q'(k)* and arrange them by increasing superscript, but keep inside each block the original generation order induced by ALG.

Thus we introduce an infinite sequence of finite sequences $\{(Y_1^k, Y_2^k, \ldots, Y_{r(k)}^k)\}_{k=1}^{\infty}$, *r(k)* = r(*k*, w) is a random variable which counts the number of sites marked ("E" or "F") in Q'(k). For every w and every *k*, Y_i^k is an element, U_{j_i}, of the sequence $\{U_j\}$, with $S(U_{j_i}) \epsilon Q'(k)$ and U_{j_i} being the i-th element in the sequence having this property. Let us consider the infinite sequence.

$$Y = (Y_1^1, Y_2^1, \ldots, Y_{r(1)}^1, Y_1^2, \ldots, Y_{r(2)}^2, Y_1^n, \ldots, \ldots, Y_{r(n)}^n, \ldots)$$

The next lemma is the crucial one to our result.

REARRANGEMENT LEMMA: *The sequence ((* $V(Y_b^a$, b = 1,. . . , r(a)), a = $1, 2, \ldots)$ *is a Bernoulli sequence of independent trials, with probability p for success.*

PROOF. The intuitive reason is that our rearrangement did not infringe or "ill-conditioned" the basic rights of any sequence element to get to this position and mark it "F" or "E" independently with the correct probability. The proof is cast in an elementary form, similar to the proof of Doob's result [Do 36] on the "futility" of gambling system, the way it is presented in Feller [Fe 68]. The situation here is much more general and we carry the proof in detail.

We have to prove that for any $s > 0$ and $z_1, \ldots, z_s \epsilon Y$

$$P(V(Z_1) . \alpha_1 \ldots A \quad V(Z_s) = \alpha_s) = \prod_{j=1} p_{\alpha_j},$$

where α_i is 0 (for "empty") or 1 (for ('full") and

$$p_{\alpha_i} = \alpha_i p + (1 - \alpha_i)(1 - p).$$

The $r(a)$'s being random variables, we have to make sure, every time we use the symbol Y_b^a, that there is indeed a b-th element in block Y", this is achieved by conditioning any event containing the symbol Y_b^a on the event $\{b \leq$ r(a)}. We introduced the following abbreviations for denoting events:

A_k^{ab} denotes the event $\{S(Y_b^a) = S(U_k)\}$

B_k^α " $\{V(U_k) = \alpha\}$

C_a^b " $\{r(a) \geq b\}$

1) *for* $s = 1$: we prove that for any a and b and for $\alpha \epsilon \{0,1\}$

$$P(*) =_{def} P(V(Y_b^a) = \alpha | b \leq r(a)) = p_\alpha.$$

The event $\{V(Y_b^a) = \alpha \wedge b \leq r(a)\}$ is the union of the disjoint events $A_k^{ab} B_b^\alpha$, $k = 1, 2, \ldots$, thus

$$P(*) = \frac{1}{P(C_a^b)} P(V(Y_b^a) = \alpha \text{ A } C_a^b)$$

$$= \frac{1}{P(C_a^b)} \sum_{\ell=1}^{\infty} P(A_k^{ab} B_k^\alpha C_a^b)$$

$$= \frac{1}{P(C_a^b)} \sum_{k=1}^{\infty} P(A_k^{ab} B_k^\alpha)$$

The last transition is due to the fact that $A_k^{ab} \subset C_a^b$ and therefore $A_k^{ab} C_a^b = A_k^{ab}$.

By definition, B_k^α depends uniquely on the outcome of the k-th trial (U_k), whereas A_k^{ab} depends only on $U_1, \ldots, U_{k-1}$, since the site where the k-th marking is executed is determined by the $k - 1$ preceeding ones. Therefore B_k^α and A_k^{ab} are independent, and

$$P(*) = \frac{1}{P(C_a^b)} P(B_k^\alpha) \sum_{k=1}^{\infty} P(A_k^{ab}).$$

The series is exactly the probability that the b-th marking in Q'(u) will eventually occur, i.e. $P(r(a) \geq b)$, and we get

$$P(*) = \frac{1}{P(C_a^b)} p_\alpha P(C_a^b) = p_\alpha.$$

2) $s = 2$: We denote

$$P(**) =_{def} P(V(Y_b^a) = \alpha \text{ A } V(Y_d^c) = \beta | b \leq r(a),\ d \leq r(c)).$$

Let $E = E(ab, cd)$ denote the event that Y_b^a preceeds Y_d^c in ALG, i.e. if $Y_b^a(\omega) = U_k(\omega)$ and $Y_d^c(\omega) = U_j(\omega)$ then $k < j$.

Clearly, if a = c, only one of the events E or $\overline{E} = E(cd, ab)$ can occur and we could assume without loss of generality that $b > d$. For the general case, we have to split $P(**)$ in two sums:

$$P(**) = \frac{1}{P(C_a^b C_c^d)}\Big[\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} P(A_k^{ab} B_k^{\alpha} C_a^b A_j^{cd} B_j^{\beta} C_c^d E) + \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} P(\ldots \overline{E})\Big].$$

Let us denote the first sum $P(**1)$ and the second $P(**2)$. Like in 1) we use the facts that $A_k^{ab} \subset C_a^b$ and $A_j^{cd} \subset C_c^d$. For j $\leq k$ $A_k^{ab} A_j^{cd} E = \emptyset$ and since for j $> k$ $A_k^{ab} A_j^{cd} \subset E$, we have for such $A_k^{ab} A_j^{cd} E = A_k^{ab} A_j^{cd}$, thus

$$P(**1) = \sum_{k=1}^{\infty}\sum_{j=k+1}^{\infty} P(A_k^{ab} B_k^{\alpha} A_j^{cd} B_j^{\beta})$$

Like above, B_j^{β} is independent of the other events, which are determined by the outcome of the j $-$ 1 first trials, thus we obtain

$$\begin{aligned} P(**1) &= P(B_j^{\beta}) \sum_{k=1}^{\infty}\sum_{j=k+1}^{\infty} P(A_k^{ab} B_k^{\alpha} A_j^{cd}) \\ &= p_\beta \sum_{k=1}^{\infty} P(A_k^{ab} B_k^{\alpha}) \sum_{j=k+1}^{\infty} P(A_j^{cd} | A_k^{ab} B_k^{\alpha}) \end{aligned}$$

Now whenever the b-th site in Q'(a) is marked and whatever its value, the d-th site in Q'(c) will be marked sooner or later, if and only if $r(c) \geq d$ and E is true. Thus for a given $A_k^{ab} B_k^{\alpha}$, the conditional probabilities of A_j^{cd} for j $> k$ add up to $P(\{r(c) \geq d\} E | C_a^b)$. Hence we get

$$\begin{aligned} P(**1) &= p_\beta \sum_{k=1}^{\infty} P(A_k^{ab} B_k^{\alpha}) P(C_c^d C_a^b E) \frac{1}{P(C_a^b)} \\ &= P(C_c^d C_a^b E) p_\alpha p_\beta, \end{aligned}$$

by part **1)** of this proof.

Similarly one proves that $(P**2) = P(C_c^d C_a^b \overline{E}) p_\alpha p_\beta$, hence

$$P(**) = \frac{1}{P(C_a^b C_c^d)} p_\alpha p_\beta [P(C_c^d C_a^b E) + P(C_c^d C_a^b \overline{E}] = p_\alpha p_\beta$$

The identities for combinations of m variables are proved by induction on their number by precisely the same arguments. We have to consider $m!$ different sums, each of which corresponds to one of the $m!$ possible permutations giving the order in which the m variables appear in the sequence $\{U_k\}$. This concludes the proof of the Lemma.

To proceed to the BB density quotient, we want to omit a finite prefix (within a finite cube) and re-index the rest of the sequence $Y_b^a(\omega)$ and

$$(C\bigcup\partial C)\bigcap(L\setminus K(\omega))=\{Y_i,\ i=1,2,3\ldots\},$$

Where $K(w) < oo$ is the smallest w such that $L\setminus K(w)$ contains marked sites coming only from C $\bigcup\partial C$, where C(w) is the infinite cluster which ALG produces in w. Now for each $n > K(\omega)$

$$C(\omega)\bigcap[Q(n)\ \setminus\ Q(K(\omega))]\ =\sum_{i=1}^{N_n(\omega)}V(Y_i) \tag{1}$$

$$\partial C(\omega)\bigcap[Q(n)\ \setminus\ Q(K(\omega))]\ =\sum_{i=1}^{N_n(\omega)}(1-V(Y_i)) \tag{1'}$$

where

$$N_n(\omega) = \text{number of sites marked in Q(n)} \setminus Q(K(\omega)). \tag{2}$$

The idea is that the perturbation of the finite cube $Q(K(\omega))$ is negligible for the infinite C and ∂C. So studying (1),(1') and their quotient is equivalent to studying $t_n(\omega)/s_n(\omega)$. Our goal is to show that there exists a set of configurations Ω" of measure **1** such that pointwise for $\omega\epsilon(\Omega)$" .

$$\exists n(\omega)\forall n>\mathrm{n(w)}\quad \frac{t_n(\omega)}{s_n(\omega)}=\frac{1-p}{p}\pm O(\frac{\log\log s_n(\omega)}{s_n(\omega)})^{\frac{1}{2}} \tag{3}$$

(The big 0 represents a constant independent of ω).

Let us first drop all those configurations w for which ALG does not find an infinite cluster. The set of such w has measure 0 and we denote its complement in Ω by R'.

The sequence $\{V(Y_i)\}_{i=1}^{\infty}$ is a sequence of Bernoulli variables, thus for large enough h(h depending on w), the h first variables of this sequence verify the strong law of large numbers with a remainder term given by the law of the iterated logarithm, i.e., noting that $E(V(Y_i))$ = p for all i, the following inequality holds with probability 1:

$$\left|\frac{\sum_{i=1}^{h}V(Y_i)}{h}-p\right|\leq(\frac{2p(1-\mathbf{p})\log\log h}{h})^{\frac{1}{2}} \tag{4}$$

Thus if we drop another O-probability set from R', denoting what remains by Ω", we have

$$\forall\omega\epsilon\Omega\text{"}\ \exists h(\omega)\forall h\quad (h > h(w) \Rightarrow \mathbf{(4)}). \tag{5}$$

Let $\omega\epsilon\Omega$" be given. This yields a finite value to the random variable. $K = K(w)$. The cube $Q(K(\omega))$ includes all the finite clusters marked by ALG and their boundaries; any marked site not belonging to $Q(K(\omega))$ is in C(w) $\bigcup \partial C(\omega)$. Denote $A_K(w) = |C(\omega) \bigcap Q(K(\omega))|$. Clearly, for n > $K(w)$

$$|C \bigcap Q(n)| = A_K(\omega) + \sum_{i=1}^{N_n(\omega)} V(Y_i) \tag{6}$$

We choose n(w) = 2 $\max(K(\omega), h(w))$. Then for n > n(w), $N_n(\omega)$ > n/2 and therefore $N_n(\omega) > h(w)$. Dividing both sides of (6) by $N_n(\omega)$ and using (4), we get

$$p - R_n \lesssim \frac{|C \bigcap Q(n)|}{N_n(\omega)} \leq \frac{A_K(\omega)}{N_n(\omega)} + p + R \tag{7}$$

where $R_n = [(2p(1-p)\log\log N_n(\omega))/N_n(\omega)]^{\frac{1}{2}}$

As n increases, $N_n(w) \Rightarrow \infty$ and $A_K(\omega)/N_n(\omega)$ is $O(1/N_n(\omega))$ which for large n is absorbed by R_n, since A, > $O(1/N_n(\omega))^{\frac{1}{2}}$. Similarly, denoting $B_K(\omega) = |\partial C(\omega) \bigcap Q(K(w))|$, we have

$$|\partial C(\omega) \bigcap Q(n)| = B_K(\omega) + \sum_{i=1}^{N_n(\omega)} (1 - V(Y_i))$$

so that for large enough n

$$1 - p - O(R_n) \leq \frac{|\partial C \bigcap Q(n)|}{N_n(\omega)} \leq 1 - p + O(R_n). \tag{8}$$

Combining (7) and (8) and setting $q = 1 - p$, $R = O(R_n)$ for the ease of description, we' get for large n

$$\frac{q-R}{p+R} \leq \frac{t_n}{s_n} \leq \frac{q+R}{p-R}. \tag{9}$$

But

$$\frac{q-R}{p+R} = \left(\frac{q-R}{p}\right)\left(\frac{1}{1+\frac{R}{p}}\right)$$

$$= \left(\frac{q-R}{p}\right)\left(1 - \frac{R}{p} + \frac{R^2}{p^2} - \cdots\right) = \frac{q}{p} - \frac{R}{p^2} + O(R^2)$$

-and

$$\frac{q+R}{p-R} = \frac{q}{p} + \frac{R}{p^2} + O(R^2).$$

As $R \to 0$ with $n \to \infty$, $O(R^2)$ may be neglected, hence

$$\left|\frac{t_n}{s_n} - \frac{1-p}{p}\right| \leq O(R) = O(R_n)$$

Knowing that $N_n(\omega), t_n(\omega)$ and $s_n(w)$ are of the same size up to a factor, this is equivalent to (3). Our proof is concluded.

6. Concluding Remarks

If the infinite cluster is unique, then the function $R_n(\omega)$, which counts how many sites are marked by ALG in Q(n), has a sharply determined positive density (and as we have noted, the proof of the BB-density-quotient is obtained by a simple application of the Ergodic Theorem [NS]. Conversely, proving or assuming certain facts about $R_n(w)$ will imply uniqueness [KS].

By taking the quotient of densities of ∂C and C the factor $R_n(w)$ is cancelled and we proved (using ALG) Theorem 1 for any cluster, without uniqueness. The importance of studying the properties of infinite clusters is that in doing simulation and experiments with finite lattices one never knows whether he measures a large finite cluster or part of an infinite cluster. This does not matter provided one proves that the asymptotics of size-n-clusters [ADS 80] coincides with that of the expanding "cubic" sections of infinite clusters. The connection with the formulation of the ergodic theorem is clear, but the theorem itself cannot be applied, because of a random behavior of $R_n(w)$ or irregularities of the lattice L, while the algorithmic method and the rearrangement Lemma do apply, here and perhaps in similar situations.

references

[ADS 80] M. Aizenmann, F. Delyon and B. Souillard, Lower Bounds on the cluster size distribution *J. of Stat. Physics 23,* 3, 267-280.

[Do 36] J.L. Doob, Note on probability, *Annals of Math.* 37 (1936), 363-367.

[Fe 69] W. Feller, *An* Introduction *to* Probability *Theory* and its Applications, Vol 1, 3rd Edition, Wiley, N.Y. 1969.

[Fi 61] N.E. Fisher, Critical probabilities for cluster size and percolation problems. *J.* Math. Phys. 2 (1961), 620-627.

[H 79] -A. Hankey, Three properties of the infinite cluster in percolation theory *J. Phys. Al 1* (1978) L49-L55.

[Hr 78] T.E. Harris, A lower bound **for** the critical probability in a a certain percolation process, *Proc. Camb.* Phil. *Soc.* 59 (1960), 13-20.

[Ke 80] The critical probability of bond percolation on a square lattice equals 1/2. *Comm.* Math. *Phys.* 74 (1980), 41-59.

[KS 81] S.T. Klein and E. Shamir. *On the* Shape of infinite *clusters in* percolation models. Unpublished manuscript 1981.

[NS 81] C.M. Newman and L.S. Schulman, Infinite clusters in percolation models. *J. Stat. Phys. 23* (1981) 267-280.

[St 79] D. Stauffer, Scaling theory of percolation clusters. *Physics Reports* 54 (1979), 1-74.

[Sh 82] E. Shamir, *How to get* an *almost sure performance of combinatorial algorithms* Stanford CS Report 1982 (to appear).