Metastability effects in bootstrap percolation

1988 J. Phys. A: Math. Gen. 21 3801

(http://iopscience.iop.org/0305-4470/21/19/017)

J. Phys. A: Math. Gen. 21 (1988) 3801-3813. Printed in the UK

Metastability effects in bootstrap percolation†

M Aizenman‡ and J L Lebowitz§

‡ Courant Institute of Mathematical Sciences and Department of Physics, New York University, 251 Mercer Street, New York, NY 10012, USA
§ Departments of Mathematics and Physics, Rutgers University, New Brunswick, NJ 08903, USA

Received 19 April 1988

Abstract. Bootstrap percolation models, or equivalently certain types of cellular automata, exhibit interesting finite-volume effects. These are studied here at a rigorous level. We find that for an initial configuration obtained by placing particles independently with probability $p \ll 1$, at sites of $\mathbb{Z}^d (d \geqslant 2)$, the density of the 'bootstrapped' (final) configurations in the sequence of cubes $(-L/2, L/2)^d$ typically undergoes an abrupt transition, as L is increased, from being close to 0 to the value 1. With L fixed at a large value, the mean final density as a function of p changes from 0 to 1 around a value which varies only slowly with L—the pertinent parameter being $\lambda = p^{1/(d-1)} \ln L$. The driving mechanism is the capture of a 'critical droplet'. This behaviour is analogous to the decay of a metastable state near a first-order phase transition, for which the present analysis offers some suggestive ideas.

1. Introduction

The phenomenon of metastability is of great interest in various situations, including numerical studies where questions sometimes arise as to whether an observed effect is a phase transition or a manifestation of metastability. Examples with such uncertainty have been encountered in studies of the ergodic properties of energy-preserving spin flip dynamics [1] and of some cellular automata models [2]. Both of these examples are related to bootstrap percolation [3], which is of interest in its own right. Our purpose here is twofold: to provide a rigorous analysis of the finite-volume effects in bootstrap percolation—including the proof of a sharp transition in suitably scaled variables, and to illustrate the role of critical droplets in the phenomenon of metastability by a particularly simple example.

Bootstrap percolation was (apparently) first considered by Chalupa *et al* [3] and was subsequently studied (rediscovered) by many others in a variety of contexts. The model is simply stated: on a d-dimensional lattice, points are independently occupied with a (low) density $1 - e^{-p}$ $(\simeq p)$ and the resulting configuration is taken as the initial state for dynamics based on some simple local rules, in which the occupation status of a point is updated according to the configuration of its neighbours. In the example to which we shall explicitly refer the rules are

(i) the subsequent updates occur simultaneously at all the lattice sites;

(ii) occupied sites remain occupied;

(iii) if an unoccupied site has (at a given time) two or more occupied neighbours, it becomes occupied (at the next update).

† Presented at the conference on *Mathematical Problems in Statistical Mechanics* held at Heriot-Watt University on 3-5 August 1987. Some other papers from the conference were published in *Journal of Physics A: Mathematical and General* 1988, volume 21, pp 1741-86.

0305-4470/88/193801+13$02.50

The primary object of interest is the final configuration of the occupation variables. In the example given above, the sites which are eventually occupied form the minimal covering of the initial set by disjoint rectangles, with separation of at least two empty lattice layers.

We say that a region Λ is internally spanned if $\Lambda \cap \mathbb{Z}^d$ is entirely covered in the final configuration for the dynamics restricted to Λ (with the analogue of the free boundary conditions), and denote the spanning probability by the function R:

$$R(L, p) = \text{Prob}(\Lambda_L \text{ is internally spanned}) \tag{1.1}$$

with $\Lambda_L \equiv (-L/2, L/2)^d$.

It has been observed [2-6] that, as p is varied, at large values of L, the spanning probability undergoes a rapid transition from 0 to 1, at values of p which are not much affected by the doubling of L. One may ask whether in the infinite volume there is an associated phase transition at a non-zero value of p (see [2]). It is now well understood that the answer to that question is negative [4-6]. As to the observed effect: our main results identify a scale—that of $\lambda = p^{1/(d-1)} \ln L$—at which there is indeed a sharp transition, defining (at least for subsequences of $p \to 0$) a critical λ.

The range of length scales with λ below its critical value is reminiscent of a metastable regime. We also find in the bootstrap percolation a vaguely defined value of p above which that metastable regime does not occur, since the required length scale no longer exceeds the correlation length of the independent model (which here is $O(1)$). The last phenomenon is reminiscent of the so-called 'spinodal decomposition', which has not yet been fully elucidated for any model with short-range dynamics.

Other variants of the basic set-up are possible, and in § 2 we present some general conditions under which our analysis is applicable. In particular, the time evolution presented above is that of the cellular automata model Q234 which has been extensively discussed by Vichniac [2]. Our own work originated from discussions with Stauffer, cf [1], and later with Griffeath and Vichniac, cf [2]. After deriving the basic results we discovered [6] where Lenormand and Zarcone essentially solved the problem—our results can be considered as a rigorisation and extension of what they analysed heuristically and observed numerically.

Let us now state our main result on the behaviour of the spanning probability $R(L, p)$. As already mentioned, it is instructive to consider it as a function of a scaled density parameter λ:

$$\lambda = p^{1/(d-1)} \ln L$$

with which we write

$$R(L, p) = Q_p(\lambda). \tag{1.2}$$

Theorem 1. (i) Any sequence of values of p convergent to 0 has a subsequence p_n for which the functions $Q_{p_n}(\cdot)$ converge to a step function:

$$Q_{p_n}(\lambda) \to \Theta(\lambda - \lambda_c). \tag{1.3}$$

The critical value λ_c may depend on the subsequence p_n (meaning that our proof has not ruled out such a possibility), but in any case obeys uniform bounds: $c_1 \leq \lambda_c \leq c_2$ with the constants c_i (>0) described by (3.8). The convergence in (1.3) is pointwise at all $\lambda \neq \lambda_c$.

(ii) More explicitly, for small p, $Q_p(\cdot)$ are approximate step functions, in the sense that for each p there is a threshold $\lambda_c(p)$ at which $Q_p(\lambda)$ changes from being close to 0, to being close to 1. The behaviour of $R(L, p)$ in the two regimes is described by

$$R(L,p)=\exp\{-p^{-1/(d-1)}[\lambda_c(p)-\lambda+o(1)]\} \qquad \text{for} \quad o(1)<\lambda<\lambda_c(p) \tag{1.4}$$

and

$$R(L,p)\geq 1-\exp(-\text{constant}\times L^{d-1}) \qquad \text{for} \quad \lambda>\lambda_c(p) \tag{1.5}$$

with $\lambda_c(p)$ obeying the bounds mentioned in (i) and o(1) representing terms which, for $p\to 0$, vanish uniformly in L.

(iii) The transition regime defined by

$$Q_p(\lambda)\in[\varepsilon, 1-\varepsilon] \tag{1.6}$$

has, for $p\to 0$, a vanishingly small width in terms of λ. Explicitly, for each p:

$$\Delta\lambda\leq[1+o(1)]\ln\varepsilon^{-1}p^{1/(d-1)} \qquad (\leq \text{constant}\times\ln\varepsilon^{-1}/\ln L). \tag{1.7}$$

While in the above theorem the spanning probability is considered as a function of L at a fixed p, it is often more natural to study its dependence on the density p at a fixed large value of L. In that case the dramatic change in R occurs at values of p which are of the order of constant/ln L. In numerical simulations these slowly varying thresholds may appear to be volume independent and suggest the existence of a phase transition in the infinite system at a finite density. In fact, the behaviour is best interpreted as a metastability phenomenon.

The connection to metastability can be seen through the proof of theorem 1. In particular, in lemma 5 we show that the mechanism by which large regions are spanned involves the occurrence of 'critical droplets' in the initial configuration. These are *local* events whose occurrence in the vicinity of a given site has a very small probability when p is small ($\approx\exp(-\text{constant}\times p^{-1/(d-1)})$). For large L, the probability of such a locally rare event occurring somewhere in Λ_L undergoes a sharp transition as a function of the parameter λ. The situation is quite analogous to metastability phenomena in the vicinity of first-order phase transitions. For example, in the ferromagnetic Ising model with Glauber dynamics [7, 8] at low temperatures (see § 5) the magnetic field h will play the role of p; for large L there will be a fairly well defined critical field beyond which the system stops being metastable. In § 5 the results described in theorem 1 are reformulated in terms of $M(L, p)$, the probability that the origin is occupied in the final configuration. In § 5 we show that $M(L, p)$ has the same asymptotic behaviour as $R(L, p)$ and then argue that this (qualitatively) should also be the behaviour of $M_T(t, p)$—the probability that the origin is occupied at time t in the dynamical system obtained by applying the rules (i)-(iii) to the infinite lattice $\mathbb{Z}^d$ in which the sites are initially occupied independently with the probability p. This leads to a conjecture about the 'metastable behaviour' in the Ising model at low temperatures.

2. General formulation of the dynamics

Let us first state the essential features of the models we discuss. They are not all equally relevant, as we comment below.

(i) The time evolution is deterministic. The most interesting direction of extending the results may be the relaxation of this feature by an extension to stochastic models, e.g. lattice systems with Glauber dynamics (see § 5).

(ii) No annihilation of particles. We denote the configuration at (integer) times t by $\eta_t = \{\eta_t(x)\}$, with $\eta_t(x) = 0, 1$ according to whether the site x is empty ($\eta = 0$) or occupied ($\eta = 1$). Condition (ii) means that $\eta_t(x)$ is a monotone function of t, for every x. That monotonicity is not strictly essential, e.g. our analysis can be adapted to a modified bootstrap model in which particles with less than two neighbours are removed. In the latter dynamics certain finite clusters (of four particles) are stable and would play the role of 'particles' in our discussion.

(iii) Attractiveness. If at the initial time a configuration η' differs from some other configuration η only by the addition of some particles, i.e. $\eta'(x) \geqslant \eta(x)$ for all x, then this property will remain true for all subsequent times (i.e. in a natural (FKG) sense η_t $(t > 0)$ is a non-decreasing function of η_0.)

(iv) Marginal stationarity of the planar interface: (a) any slab of empty sites, parallel to a principal axis of the (cubic) lattice, of thickness greater than some D will remain empty for all times, yet (b) the adsorption of a finite cluster of particles (in the present formulation one is enough) on a stationary interface will eventually lead to a shift of the whole interface into the empty region.

Conditions (iii) and (iv) are natural for the problem we are considering, where the fully occupied configuration corresponds to the stable state and the empty configuration to an unstable stationary state.

It follows from (iii) and (iv) that a single rectangular set (finite or infinite) of occupied sites is stationary; rectangular sets being for us parallelograms aligned with the principal lattice planes. When the model's interaction also satisfies a finite-range condition, then any configuration consisting of a union of well separated rectangular clusters is stationary. In the bootstrap percolation model the converse is also true.

(v) Characterisation of stationarity. A configuration is stationary if and only if its particles form a collection of rectangular sets separated by distances greater than D.

The above assumptions are all satisfied for the bootstrap percolation discussed in § 1 for which $D = 2$. Another variant of the model, with all the listed properties and $D = 1$, is obtained by changing the rule so that a vacant site is occupied only if it has two occupied neighbours which are not on its opposite sides. Models with other values of D may be found in [2].

Condition (v) leads directly to the following deterministic lemma which plays an important role in our argument.

Lemma 1. For all $k \geqslant 1$, a necessary condition for a region Λ_L with $L \geqslant k$ to be internally spanned (in a configuration η) is that it contain at least one rectangular region whose maximal side length is in the interval $[k, 2k + D]$ which is also internally spanned.

It should be noted that the clusters spanned by bootstrap percolation depend quite delicately on the relative placement of widely separated connected clusters (see, for example, figures in [2, 6, 9]). In particular, the larger region Λ_L may be internally spanned without there being any region with length k (>2) which is internally spanned.

Proof. The configuration spanned by the restriction of η to Λ_L can be determined by the following algorithmic construction which, at the nth step, deals with a collection, $\mathscr{C}_n$, of possibly overlapping rectangular regions, each of which is internally spanned. Starting with $\mathscr{C}_1$ as the collection of all the internally spanned rectangles of an arbitrarily chosen size $k_0 \leqslant k$, we proceed by choosing at each step a pair of rectangles in the configuration whose distance is not greater than D and replacing them with the minimal rectangular region which contains both. The lemma follows by the observation that the maximal length of a rectangle in $\mathscr{C}$ increases in this process no faster than $m \to 2m+D$, and that it does reach L if Λ_L is internally spanned.

3. Critical droplets

In the proof of the main result (stated in § 1) we arrive at a more complete description of the spanning probability $R(L, p)$, which exhibits different behaviour over different length scales. In the following description we take p as fixed at a small value.

(i) For small L, $R(L, p)$—which is a polynomial in p of degree not greater than $|\Lambda_L|$ ($\approx L^d$)—decreases rapidly with L. While we do not quite prove monotonicity, we show that for p small enough the maximal value of $R(L, p)$ over lengths up to $O(p^{-1/(d-1)})$ occurs at $L=1$. For lengths of the order of $O(p^{-1/(d-1)})$, $R(L, p)$ is so small that the quantity $\sigma(L, p)$ defined by

$$R(L, p) \equiv \exp(-\sigma(L, p)p^{-1/(d-1)}) \qquad (3.1)$$

is of order $O(1)$.

(ii) After the initial transient regime, there is a plateau extending over the wide range of scales: $p^{-1}/\varepsilon \leq L \leq \exp(\varepsilon p^{-1/(d-1)})$, where the function $\sigma(L, p)$ (which determines $R(L, p)$) is approximately constant, $\hat{\sigma}(p) = O(1)$. ε is interpreted here as a small quantity which is taken to 0 at a rate slower than any (fractional) power of p. (The arguments we use suggest that the plateau starts already at $L \approx p^{-1/(d-1)}/\varepsilon$.)

(iii) When the length gets to be of the order of $L = \exp(\lambda p^{-1/(d-1)})$, with $\lambda = O(1)$, σ drops down linearly in $\lambda (= p^{1/(d-1)} \ln L)$ as

$$\sigma(L, p) = \hat{\sigma} - \lambda. \qquad (3.2)$$

At the critical value of $\lambda (= \hat{\sigma}(p))$ $R(\cdot, p)$ ceases to be small.

(iv) For values of λ beyond λ_c the spanning probability R is approximately 1, with deviations vanishing at the rate described in (1.5).

The interpretation of the behaviour of $R(\cdot, p)$ in the regimes (2) and (3) is that the bootstrap percolation is dominated by the occurrence (somewhere in Λ_L) of a fairly local 'bottleneck event', whose probability is close to $\hat{R}(p) \equiv \inf\{R(n, p) | n \geq 1\}$. The following statement establishes the order of magnitude of $\hat{R}(p)$.

Remark. While both L and n will typically denote lengths, we tend to use n for sizes which are small on what will eventually be regarded as the 'macroscopic' scale. (Nevertheless, the formulae we write hold for all lengths.)

Lemma 2. For all n and p, the quantity $\sigma(\cdot, \cdot)$ defined by (3.1) satisfies

$$D^{d/(d-1)} g(n(Dp)^{1/(d-1)}) \leq \sigma(n, p) \leq d \int_0^\infty \mathrm{d}z\, g(z)/z \qquad (<\infty) \qquad (3.3)$$

where $g(\cdot)$ is the function

$$g(z) = z \ln[1 - \exp(-z^{d-1})]^{-1}.$$

Proof. (i) The upper bound in (3.3), which expresses a lower bound on $R(n, p)$, is a consequence of the fundamental observation [4, 5] that there is an event of non-vanishing probability which ensures that all the cubes centred at the origin (or at any other given point) are internally spanned. The condition is (by property (iv)(b) of § 2) that for each $k \geq 0$ there is at least one particle on each of the $2d$ faces of the cube of length $(2k+1)$ centred at the origin. This yields

$$\begin{aligned} R(n,p) &\geq \prod_{k=0}^{n/2} \{1-\exp[-p(2k+1)^{d-1}]\}^{2d} \\ &= \exp\left(-2d \sum_{k=0}^{n/2} \ln\{1-\exp[-p(2k+1)^{d-1}]\}^{-1}\right) \\ &\geq \exp\left(-d \int_0^{n+1} \mathrm{d}s \ln[1-\exp(-ps^{d-1})]^{-1}\right) \\ &\geq \exp\left(-d p^{-1/(d-1)} \int_0^{\infty} \mathrm{d}z\, g(z)/z\right). \end{aligned} \tag{3.4}$$

(ii) For an upper bound on $R(n,p)$ we use the fact that (by property (iv)(a)) a necessary condition for a region of width n to be internally spanned is that in its partition into n/D disjoint slabs, of width D, parallel to one of the principal lattice planes, each slab contains at least one particle. Hence, for each $n \geq 1$ (not necessarily a multiple of D)

$$\begin{aligned} R(n,p) &\leq [1-\exp(-pDn^{d-1})]^{n/D} \\ &= \exp\{-(n/D)\ln[1-\exp(-pDn^{d-1})]^{-1}\} \\ &= \exp\{-D^{d/(d-1)}p^{-1/(d-1)}g[n(pD)^{1/(d-1)}]\}. \end{aligned} \tag{3.5}$$

When expressed in terms of $\sigma(n,p)$, the bounds (3.4) and (3.5) yield (3.3).

We shall now use the left-inequality in (3.3) to prove the statements made above about the transient regime (1).

Lemma 3. (i) There is $\bar{C} < \infty$ (depending only on d) such that, for each p,

$$\max\{R(n,p) | 1 \leq n \leq \bar{C}(pD)^{-1/(d-1)}\} \leq pD^d. \tag{3.6}$$

Furthermore, for p small enough the maximum in (3.6) is at $n=1$.

Proof. Since pD^d is an upper bound on the probability of there being at least one particle in Λ_D, it certainly bounds $R(n,p)$ for $n \in [1, D]$. Now let $\bar{C}$ be the first local maximum of the function $g(\cdot)$ on $[0,\infty)$. (That is a well defined number since $g(\cdot)$ is continuous, vanishes at $z=0$ and $+\infty$, and is positive elsewhere.) By the monotonicity of $g(\cdot)$ in $[0,\bar{C}]$, the lowest value the bound (3.5) takes within the region $D \leq n \leq \bar{C}(pD)^{-1/(d-1)}$ is at $n = D$, where it yields pD^d. That proves (3.6).

To see that for p small enough the minimum in (3.6) is at $n=1$, one may split the range of values of n into the three cases: $n=1$, $2 \leq n \leq 2D$, $2D \leq n \leq \bar{C}(pD)^{-1/(d-1)}$. For the last two cases $R(n,p) \leq bp^2$, with a finite constant b. (For the second regime, that follows from the necessity to have at least two particles, and for the third regime by an adaptation of the argument used above.) Thus, for $p < b$ the minimum is at $n=1$, where $R(1,p)=p$.

It should also be noted that the claim made in (1) in relation to (3.1) is an explicit consequence of the lower bound in (3.3). Our next step is to establish the existence of the plateau described in (2).

Lemma 4. For each p, and pair of positive numbers $A \gg 1 \gg B$, such that $A > 3 \ln p^{-1}$, $\sigma(\cdot, p)$ as a function of n is approximately constant throughout the regime

$$Ap^{-1} \leqslant n \leqslant \exp(Bp^{-1/(d-1)})$$

satisfying there

$$|\sigma(n,p) - \hat{\sigma}| \leqslant 2dB + 2dp^{-(d-2)/(d-1)} \exp(-A/3) \tag{3.7}$$

with some $\hat{\sigma} = \hat{\sigma}(p)$, which satisfies

$$C_1 \equiv D^{d/(d-1)} \sup\{g(z) | z \geqslant 0\} \leqslant \hat{\sigma} \leqslant d \int_0^\infty \mathrm{d}z\, g(z)/z \equiv C_2. \tag{3.8}$$

Proof. Taking as the 'seed' for the construction which led to (3.4) not a single particle at the origin but, instead, an internally spanned region Λ_k with any $k \leqslant n$, we get

$$\begin{aligned} R(n,p) &\geqslant R(k,p) \exp\left(-dp^{-1/(d-1)} \int_{kp^{1/(d-1)}}^\infty \mathrm{d}z\, g(z)/z\right) \\ &\geqslant R(k,p) \exp[-p^{-1/(d-1)} \exp(-A^{d-1}/2)] \end{aligned} \tag{3.9}$$

where in the second step it is assumed that $kp^{1/(d-1)} \geqslant A$.

For an opposite bound on the ratio of the two spanning probabilities, we note that, by the deterministic observation made in lemma 1,

$$R(n,p) \leqslant k^d n^d R_1(k,p) \tag{3.10}$$

where $R_1(k,p)$ is the maximum of the spanning probabilities of rectangular regions whose longest side length falls in the interval $[(k-D)/2, k]$. The factor $k^d n^d$ on the right-hand side of (3.10) counts the number of such rectangles in Λ_n and the bound is on the probability that at least one such rectangle is internally spanned. In order to replace $R_1(k,p)$ by $R(k,p)$, we note that, by a similar construction to the one leading to (3.4),

$$\begin{aligned} R(k,p) &\geqslant R_1(k,p) \exp\{-kd \ln[1-\exp(-p(k-D)/2)]^{-1}\} \\ &\geqslant R_1(k,p) \exp[-2dp^{-1} \exp(-A/3)] \end{aligned} \tag{3.11}$$

where we used the fact that $kp \geqslant A \gg 1$. Combining (3.11) with (3.10) we have

$$R(n,p) \leqslant k^d n^d R(k,p) \exp[2dp^{-1} \exp(-A/3)]. \tag{3.12}$$

(An observant reader may note that for $d > 2$ the power of p in the last two bounds may not be optimal. That is due to the limitation of our 'worst case' scenario for the spanning of Λ_L.)

The inequalities (3.9) and (3.12) directly imply the claim made in (3.7), with the bounds (3.8) on $\hat{\sigma}$ being deduced from those of lemma 2. In this argument, the role of the assumptions involving B is to ensure that the entropy-related factor $k^d n^d$, of (3.12), has only a negligible effect on $\sigma(n,p)$.

We now turn to justify the 'critical droplet' picture to which reference has already been made a number of times. The main idea is expressed in lemma 5, where the event that Λ_L is spanned is shown to be effectively equivalent to the occurrence in Λ_L of a localised 'bottleneck' event, whose main parameters are (i) its probability (close to $\hat{R}(p)$) and (ii) the (linear) size of its localisation, $n(p)$. Because of the existence of the plateau discussed above, and the low value of $\hat{R}(p)$, we have the freedom to choose $n(p)$ within a large range of scales without significantly affecting either the probability or the resulting entropy factor (as can be seen in (4.1) below). For convenience, we choose

$$n(p) = p^{-3}. \tag{3.13}$$

In a certain sense this value for n is much too large (with a more natural choice being $p^{-1/(d-1)}/o(1)$). However, it enables one to derive the property presented in the following lemma by means of very crude estimates.

Lemma 5. Let $A_i(L)$, $i = 1, 2, 3$, denote the following three families of events:

$A_1(L) = \Lambda_L$ is internally spanned,

$A_2(L) = \Lambda_L$ contains a translate of $\Lambda_{n(p)}$ which is internally spanned,

$A_3(L) = \Lambda_L$ contains an internally spanned block which is a translate of $\Lambda_{n(p)}$ and is centred at a point of the sublattice

$$\mathbb{Z}_{n/2^d} \equiv [n(p)/2]\mathbb{Z}^d.$$

Then, for $p \to 0$ and $L \in (n(p), \exp(p^2/2))$, these events are 'almost equivalent', in the sense that their symmetric differences carry only very small probabilities:

$$\text{Prob}(A_i(L) \Delta A_j(L)) \leq \exp(-Cp^{-2}) \tag{3.14}$$

with some $C < \infty$, independent of p (and $A \Delta B \equiv (A \backslash B) \cup (B \backslash A)$).

Remarks. (i) The event A_2 may be regarded as the occurrence of a critical droplet within Λ_L. The somewhat artificially restricted event A_3 has a simpler structure, which is close to that of a union of a family of independent local events.

(ii) The upper restriction on L in the statement of the lemma is only 'temporary', since our results will imply that, for greater L, the probabilities of all A_i are much closer to 1 than $\exp(-Cp^{-2})$.

Proof. We prove (3.14) by showing that there is an event G, whose probability is close to 1, conditioned on which the events A_i are all equivalent. We choose for G the event that each linear segment of length $n(p)/2$, aligned with one of the principal axes, has at least one occupied neighbouring site on each of its $2d$ long sides. The probability for the condition to fail on a given segment is clearly bounded by $2d \exp(-pn(p)/2)$ and the number of such linear segments does not exceed dL^d. Hence

$$\text{Prob}(G^c) \leq 2d^2 L^d \exp(-pn(p)/2). \tag{3.15}$$

On the condition that G occurs, the event A_2 (and hence also A_3) is easily seen to yield A_1. For the converse direction we make use of lemma 1, by which A_1 implies that there is in Λ_L an internally spanned region whose diameter is between $[n(p) - D]/2$ and $n(p)$. Any such region is entirely contained in at least one of the translates of Λ_n by vectors in the lattice $\mathbb{Z}_{n/2^d}$. Under the condition G, that translate is also internally spanned. Therefore

$$A_i(L) \Delta A_j(L) \subset G^c \tag{3.16}$$

for $i, j = 1, \ldots, 3$ and the claim (3.14) follows from the bound (3.15), with a simple choice of C.

4. The spanning probability $R(L, p)$

Using the results of § 3, we shall now prove the main theorem, which is stated in § 1.

Proof of theorem 1. The behaviour of the spanning probability $R(L, p)$ in the intermediate regimes is derived in lemmas 3 and 4. By lemma 5, for L in the range $(n(p), \exp(p^2/2))$ $R(L, p) = \text{Prob}(A_1(L))$ is well approximated by the probability of the 'critical droplet' event $A_3(L)$. That event is easy to estimate, since it concerns a fixed number of translates (corresponding to the 2^d different sublattices of $\mathbb{Z}_{n/2^d}$) of an event which is just an intersection of independent local events. The obvious arguments yield

$$2^d[L/n]^d R(n, p) \geq \text{Prob}(A_3(L)) \geq 1 - [1 - R(n, p)]^{[L/n]^d}$$
$$\approx 1 - \exp\{-R(n, p)[L/n]^d\} \qquad (4.1)$$

with $n = n(p)$.

Combining the above with (3.14), and the estimate provided by lemma 4 for $R(n, p)$, we get the following.

(i) For L with $p^{1/(d-1)} \ln L \equiv \lambda < \hat{\sigma}(p)$: $R(L, p) = o(1)$, satisfying

$$R(L, p) = \exp\{-p^{-1/(d-1)}[(\hat{\sigma}(p) - \lambda) + o(1)]\} \qquad (4.2)$$

as claimed in (1.4).

(ii) For L $(\leq \exp(p^2/2))$ with $\lambda > \hat{\sigma}(p)$: $R(L, p) = 1 - o(1)$, with our initial bound being

$$R(L, p) \geq 1 - \exp\{-p^{-1/(d-1)}[(\lambda - \hat{\sigma}(p)) + o(1)]\}. \qquad (4.3)$$

The above bound is not optimal and can be improved by the following renormalisation argument which also extends the result beyond the restriction $L \leq \exp(p^2/2)$ made earlier.

Let us partition the original lattice $\mathbb{Z}^d$ into a similar lattice of blocks of size Λ_{L_0}, with some $L_0 \geq 1$. For each initial particle configuration we generate a configuration on the rescaled lattice by regarding a site as occupied if the corresponding block is internally spanned, and then performing on this system a bootstrap percolation process with $D = 1$ (i.e. one where only direct neighbours have a combined effect, see § 2). It then follows that a sufficient condition for Λ_{kL_0} on the original lattice to be internally spanned is that the corresponding region Λ_k on the rescaled lattice is spanned by its process. Thus, $R(\cdot, \cdot)$ obeys the following 'renormalisation inequality'

$$R(L, p) \geq \tilde{R}(L/L_0, R(L_0, p)) \qquad (4.4)$$

where $\tilde{R}$ is the function R for the particular bootstrap percolation process with $D = 1$, and (for simplicity) we assume that L is a multiple of L_0.

When p is close enough to 1, the standard contour expansion for the site percolation problem (related to the Peierls expansion) converges and proves that

$$\tilde{R}(k, p) \geq 1 - \exp(-\text{constant} \times \alpha k^{d-1}) \qquad (4.5)$$

for $p \geq 1 - \exp(-\alpha)$ with sufficiently large α.

To improve on (4.3), we choose p_0 to be any density high enough so that (4.5) is satisfied, e.g. any point above the Peierls value (presumably, any point above the actual percolation threshold will also do) and let L_0 be the smallest L for which $R(L, p) \geq p_0$. (The value of λ corresponding to L_0 is of the order of $\hat{\sigma}(p) + o(1)$.) The combination of the renormalisation inequality (4.4) with the high-density estimate (4.5) yield the improved bound (1.5).

(iii) Let us now turn to the transition regime, which may naturally be defined by the condition (1.6) with some arbitrary (small) ε, which will be kept fixed as $p \to 0$. The basic inequalities (3.14) and (4.1) imply that, throughout this regime,

$$\ln\left(\frac{1}{\varepsilon - o(1)}\right) \leqslant [L/n(p)]^d R(n, p) \leqslant [\varepsilon + o(1)]2^d \tag{4.6}$$

with o(1) representing ε-independent quantities which vanish when $p \to 0$. Consequently, when L varies throughout the transition regime defined by (1.6), at a fixed p, the variation of λ (around the value $\hat{\sigma}(p)$) is bounded by

$$\Delta\lambda \leqslant [1 + o(1)] \ln(\varepsilon^{-1}) p^{1/(d-1)} \leqslant \frac{\text{constant} \times \ln \varepsilon^{-1}}{\ln L}. \tag{4.7}$$

That establishes (1.7).

The results derived above are stated as (ii) and (iii) of this theorem. Part (i) is implied by standard compactness arguments.

Without stating this as a theorem, let us point out that the most drastic manifestation of the transition discussed here may be seen by observing the density of the bootstrapped configuration in an increasing sequence of cubes with consistent initial configurations (obtained either by randomly extending the configuration at each step into $\Lambda_{L+1} \backslash \Lambda_L$, or equivalently by generating the initial configuration directly for all $\mathbb{Z}^d$). Our analysis can be applied to show that typically, for $p \ll 1$, in such a process the density of the occupied sites jumps at a certain random value of L from being of the order of $O(p)$ to 1. (Typically, the value of λ corresponding to that random L deviates from λ_c by not more than the order of $\Delta\lambda$ discussed above.)

5. The time evolution picture

5.1. Behaviour of the time-dependent density function

So far our discussion centred on the quantity $R(L, p)$ which describes properties of the bootstrap percolation in the limit where $L \to \infty$ after first taking the time to infinity. It is also very instructive to consider the other order of limits, namely consider the infinite system's time evolution. The basic information on that process is conveyed by the quantity

$$M_T(t, p) = \text{Prob}(\eta_t(0) = 1) \tag{5.1}$$

(in the notation of § 2), which we view as analogous to the time-dependent magnetisation for Glauber dynamics of a spin system.

We conjecture here that $M_T(t, p)$ also behaves like an approximate step function (somewhat like $R(\cdot, \cdot)$ when $p \to 0$) in terms of the parameter $\lambda_T = p^{1/(d-1)} \ln t$ in the sense that, for some critical value $\lambda_{T,c}$:

$$\lim_{\substack{p \to 0, t \to \infty \\ p^{1/(d-1)} \ln t \to \lambda}} M_T(t, p) = \begin{cases} 0 & \text{for} \quad \lambda < \lambda_{T,c} \\ 1 & \text{for} \quad \lambda > \lambda_{T,c}. \end{cases} \tag{5.2}$$

This statement may, of course, be rephrased in a stronger way, since the convergence of M to either 0 or 1 means also that the probability measure describing the entire configuration η_t converges to a limit, which in one case is concentrated on the empty configuration and in the other case on the completely filled one (with $\eta(\cdot) \equiv 1$). For reasons mentioned in the introduction (and in the abstract) we regard (5.2) as a demonstration of a metastability phenomenon, even though (5.2) in itself does not convey the full information about the demise of the metastable phase, e.g. at the level of detail reached in the case studied by Cassandro *et al* [10]. (We thank R Schonmann for calling our attention to this difference.)

A related quantity, which is somewhat easier to analyse, is

$$M(L, p) = \text{Prob}(0 \text{ is occupied in the final configuration for the bootstrap percolation in } \Lambda_L). \tag{5.3}$$

It may be noted that $M_T(t, p)$ and $M(L, p)$ are both monotone increasing functions in each of their arguments (unlike $R(L, p)$) and they clearly satisfy

$$M(t^{1/d}/2, p) \leqslant M_T(t, p) \leqslant M(t, p). \tag{5.4}$$

For $M(\cdot, \cdot)$ we have the following result.

Theorem 2.

$$\lim_{p\to 0} \sup_L \{|M(L, p) - R(L, p)|\} = 0. \tag{5.5}$$

In particular, for $p \ll 1$, $M(L, p)$ is an approximate step function in the variable $\lambda = p^{1/(d-1)} \ln L$, with the transition from 0 to 1 occurring at $\lambda_c(p)$ (of theorem 1).

Proof. We approach $M(L, p)$ by writing it as follows (with $n < L$):

$$R(L, p) \leqslant M(L, p) = [M(L, p) - M(n, p)] + M(n, p) \tag{5.6}$$

where the inequality expresses the obvious fact that if Λ_L is internally spanned then 0 is occupied at the final configuration of Λ_L. For a bound on the right side, which will of course be of interest only for L with λ below λ_c, we note that

$$\begin{aligned} M(L, p) - M(n, p) &= \text{Prob}(0 \text{ is covered in the final configuration of } \Lambda_L \text{, but not in the final configuration of } \Lambda_n) \\ &\leqslant \text{Prob(the final configuration of } \Lambda_L \text{ includes a cluster which connects } 0 \text{ with } \Lambda_n^c). \end{aligned} \tag{5.7}$$

Under the condition G discussed in the proof of lemma 5, if the event on the right-hand side of (5.7) occurs, then Λ_L is spanned. Hence

$$M(L, p) - M(n, p) \leqslant R(L, p) + \text{Prob}(G^c). \tag{5.8}$$

With $n = n(p)$ of (3.13), we have $\text{Prob}(G^c) \to 0$ as $p \to 0$ (by (3.15)). Therefore it only remains to show that

$$\lim_{p\to 0} \sup\{M(n, p) | n \leqslant 1/p^3\} = 0 \tag{5.9}$$

which makes sense since $1/p^3$ is far below $\exp(\lambda_c p^{-1/(d-1)})$ where the transition of $R(L, p)$ occurs.

By the arguments of (5.7) and (3.5) we get

$$
\begin{aligned}
M(n+1,p)-M(n,p) \\
&\leqslant \text{Prob(in the final configuration of } \Lambda_{n+1}\ 0 \text{ is connected to } \partial\Lambda_n) \\
&\leqslant 2d\{1-\exp[-pD(\tfrac{1}{2}n)^{d-1}]\}^{n/2D} \qquad (5.10)
\end{aligned}
$$

which easily implies (5.9) (e.g. see the discussion in the proof of lemma 3).

We note that, in view of the relations (5.4), part of the conjectured behaviour (5.2) does follow from theorem 2, namely one can produce values $\lambda_{T,-}$ and $\lambda_{T,+}$ (using the bounds provided for λ_c by (3.8)) with which

$$
\lim_{\substack{p\to 0, t\to\infty \\ p^{1/(d-1)}\ln t\to\lambda}} M_T(t,p)=\begin{cases}0 & \text{for } \lambda<\lambda_{T,-} \\ 1 & \text{for } \lambda>\lambda_{T,+}.\end{cases} \qquad (5.11)
$$

5.2. A conjecture for Glauber dynamics

The bootstrap percolation problem discussed here is related by strong analogies to the question of the metastable behaviour in systems undergoing first-order phase transitions. For example, at low temperatures the infinite-volume Ising system, say with nearest-neighbour ferromagnetic interaction J, undergoes a first-order phase transition at magnetic field $h=0$, with the density (of up spins) changing discontinuously from $\rho_-\approx 0$ for $h=0^-$ to $\rho_+\approx 1$ for $h=0^+$. At $h=0$ there are two phases, which can be obtained by taking the infinite-volume limit of Gibbs states in finite regions $\Lambda\subset\mathbb{Z}^d$ with '$-$' or '$+$' boundary conditions. For $h\neq 0$, on the other hand, there is only one phase and the infinite-volume limit of the Gibbs measures $\mu(\eta;\, h, \Lambda, \pm)$ is independent of the boundary conditions. Under the Glauber dynamics at $h\neq 0$ any initial state converges to the unique Gibbs state. Nevertheless, at low temperatures and small $h>0$, starting from a typical '$-$' configuration one finds that for long times the state of the system closely resembles the '$-$' phase. Locally, the change from this 'metastable state' appears (experimentally and in computer simulations) to occur rather abruptly on the timescale of the 'incubation' time.

A number of attempts have been made to develop a conceptually clear, and computationally useful, notion of the metastable state and of its demise. Without entering into this discussion, let us conjecture here that the time-dependent density in the spin system described above satisfies the suitable analogue of (5.2) with p replaced by the magnetic field h, and the convergence of M to 0 or 1 replaced by the convergence of the measure (describing the spin configuration at the time t) to the '$-$' phase in one case and the '$+$' phase in the other.

Some results which may be useful for the analysis of the above conjecture may be found in the work of Capocaccia *et al* [8] on the 'restricted ensemble' of Penrose and Lebowitz [7].

Acknowledgments

We wish to thank P Leath for introducing us to bootstrap percolation and S Goldstein for an early (unpublished) explanation of the basic property of the infinite system. It is also a pleasure to thank D Stauffer for extensive discussions of a work which led us to consider the problem studied here and D Griffeath, G Vichniac, O Penrose and R Schonmann for other stimulating discussions. MA gratefully thanks O Penrose and J M Ball for their kind hospitality at the Heriot-Watt University, where some of this work was done.

MA was supported in part by NSF grant no PHY-8605164. JLL was supported in part by NSF grant no DMR 86-12369.

References

[1] Herrmann H J, Carmesin H O and Stauffer D 1987 *J. Phys. A: Math. Gen.* **20** 4939
[2] Vichniac G 1986 *Disordered Systems and Biological Organization* (*NATO ASI Series*) vol F20, ed E Bienenstock (Berlin: Springer)
[3] Chalupa J, Leath P L and Reich G R 1979 *J. Phys. C: Solid State Phys.* **12** L31
[4] Goldstein S 1981 private communication
[5] van Enter A 1987 *J. Stat. Phys.* **48** 943
[6] LeNormand R and Zarcone C 1984 *Kinetics of Aggregation and Gelation* ed F Family and D P Landau (Amsterdam: Elsevier)
[7] Penrose O and Lebowitz J L 1971 *J. Stat. Phys.* **3** 211
[8] Capocaccia D, Cassandro M and Olivieri E 1974 *Commun. Math. Phys.* **39** 185
[9] Griffeath D 1986 *Pictures of critical configurations* unpublished (available upon request from University of Wisconsin)
[10] Cassandro M, Galves A, Olivieri E and Vares M E 1984 *J. Stat. Phys.* **35** 603