The Annals of Probability
2000, Vol. 28, No. 4, 1528–1562

INTERNAL DLA AND THE STEFAN PROBLEM

BY JANKO GRAVNER[1] AND JEREMY QUASTEL[2]

University of California and University of Toronto

Generalized internal diffusion limited aggregation is a stochastic growth model on the lattice in which a finite number of sites act as Poisson sources of particles which then perform symmetric random walks with an attractive zero-range interaction until they reach the first site which has been visited by fewer than α particles, at which point they stop. Sites on which particles are frozen constitute the *occupied set*. We prove that in appropriate regimes the particle density has a hydrodynamic limit which is the one-phase Stefan problem. This is then used to study the asymptotic behavior of the occupied set. In two dimensions when the walks are independent with one source at the origin and $\alpha = 1$, we obtain in particular that the occupied set is asymptotically a disc of radius $K\sqrt{t}$, where K is the solution of $\exp(-K^2/4) = \pi K^2$, settling a conjecture of Lawler, Bramson and Griffeath.

0. Introduction. Internal diffusion limited aggregation is a stochastic growth model in which a set in the integer lattice grows by addition of boundary sites hit by random walks produced within the set. The model also goes by the names anti-DLA and diffusion limited erosion. It has been used in physics as a model of erosion processes, but its introduction into the mathematical literature was from the unlikely direction of an algebraic construction. Diaconis and Fulton [7] consider a vector space whose basis elements are finite subsets of the lattice, and introduce a *smash product* making it a commutative algebra. The product $\{x\} * A$ of a singleton $\{x\}$ with a set A containing it is given by $\sum_y p(x, y)(A \cup \{y\})$, where $p(x, y)$ is the probability that y is the first site in the complement of A visited by a symmetric nearest-neighbor random walk starting at x. For general finite sets A and B, the smash product is $A * B = \{x_n\} * \{x_{n-1}\} * \cdots * \{x_1\} * (A \cup B)$, where $A \cap B = \{x_1, \dots, x_n\}$, the key observation being that this is independent of the ordering $\{x_1, \dots, x_n\}$. It can now be extended to general vectors by linearity and gives a well-defined product. The construction admits many generalizations, but already in this simple case there is the very interesting question of the behavior of the.random set A_n obtained from $A_0 = \{0\}$ by successively adding on the first point y in the complement visited by a symmetric nearest-neighbor random walk starting at the origin.

Received March 1998; revised January 2000.

[1]Supported in part by NSF Grant DMS-97-03923 and Republic of Slovenia's Ministry of Science Grant J1-8542-0101-97.

[2]Supported in part by NSF Grant DMS-95-04791, NSERC Grant 203087-98 and Alfred P. Sloan Foundation Fellowship.

AMS 1991 *subject classifications*. Primary 60K35; secondary 82A05.

Key words and phrases. Interacting particle system, free boundary problem, growth model, shape theory.

In [3], Bramson, Griffeath and Lawler studied the asymptotic shape of A_n and found that in any dimension d, A_n is asymptotically a ball $B(0,r)$ of radius $r = a_d n^{1/d}$ around the origin, where $a_d = (d\Gamma(d/2)/2)^{1/d}/\sqrt{\pi}$, so that $B(0,r)$ has volume n. More precisely, they proved that, for every $\delta > 0$, with probability 1, for sufficiently large n,

$$B\big(0,(1-\delta)a_d n^{1/d}\big) \subset A_n \subset B\big(0,\big(1+\delta\big)a_d n^{1/d}\big).$$

In another version of the model, particles are produced at the origin as a Poisson process at rate 1 and perform continuous-time random walks jumping at rate 1 to each nearest neighbor, until they hit the boundary, at which time the hitting site is added to the set. Now one has a set A_t and a configuration of particles on it together performing a continuous-time Markov process. In dimensions $d \geq 3$, [3] showed that the set has an asymptotic shape which is still a ball of radius $a_d t^{1/d}$ around the origin. Their proof relies on the fact that $t^{1/d}$ grows slowly enough in $d \geq 3$ that particles essentially hit the boundary before the next particle is produced, and therefore the model can be reduced to the discrete-time case.

In dimensions 1 and 2, however, the methods of [3] break down, and in fact, computer simulations show a nontrivial density of particles in the occupied region. They conjectured that in two dimensions the set grows as $K\sqrt{t}$ with a constant K which could, in principle, be computed with the following heuristic argument. Assume that the occupied set A_t is asymptotically a ball $B(0, K\sqrt{t})$ around the origin. The expected number of particles that hit the boundary before time T is then, after Brownian scaling,

$$T\int_0^1 P_0\bigg(\sup_{0\leq s\leq 1-t}(t+s)^{-1/2}|\sqrt{2}\beta_s| \geq K\bigg)dt, \tag{0.1}$$

where β_s is a standard Brownian motion starting at the origin. (We use the convention that random walks jump at rate 1 to each nearest neighbor and therefore the diffusion approximation is $\sqrt{2}\beta_s$.) On the other hand, the number of particles which hit by time T is clearly

$$T\pi K^2. \tag{0.2}$$

Bramson, Griffeath and Lawler conjectured that the correct growth rate is the unique K for which these expressions coincide.

In this article we prove the conjectures of [3] by showing that internal DLA has a hydrodynamic limit which is the one-phase Stefan problem. This is then used to study the asymptotic shape of the occupied set. For example, in the two-dimensional case described above, the occupied set grows as $K\sqrt{t}$, where

$$\exp(-K^2/4) = \pi K^2,$$

which one can check is the unique K for which (0.1) and (0.2) coincide by solving the heat equation in two dimensions with a source at the origin and Dirichlet boundary conditions on an expanding circle of size $\sqrt{t}$.

Our approach yields somewhat weaker shape theorems than the methods of [3] but has the advantage of being rather robust: it works in all dimensions and extends readily to interacting versions of the model. In particular, we consider particles produced at rates of order ε^{-d} at various sites in $\mathbb{R}^d$, then performing interacting random walks on the reduced lattice $\varepsilon\mathbb{Z}^d$ at rates of order ε^{-2}. The interaction is zero-range: the jump rate of a particle depends on the number of particles at that site. We also choose some positive integer α and decree that the first α particles at a site are frozen. The set of frozen sites is the occupied set $A_\varepsilon(t)$ and we show that it is asymptotically a ball of radius of order

$$\begin{cases} t^{1/d}, & d \geq 2, \\ (t \log t)^{1/2}, & d = 1. \end{cases}$$

We comment briefly on the relation between our scaling and that in [3]. If we consider time $t = 1$ in our model and some small $\varepsilon > 0$, then we will have created $O(\varepsilon^{-d})$ particles which have each taken $O(\varepsilon^{-2})$ steps of size ε, producing an occupied set which is a ball of order 1. This corresponds to taking times of order ε^{-2} in [3], at which time-order ε^{-2} particles will have been created and the occupied set will be a ball of radius $O(1)$ only if we rescale the lattice width to be $O(\varepsilon^{2/d})$. In $d \geq 3$, therefore, our model is in some sense a renormalized version of the model in [3], with each of their particles corresponding to ε^{2-d} of our particles. The surprise is that the size of the occupied set is the same in both models, which means that the dynamics is essentially unaffected by this renormalization.

The convergence to the Stefan problem is proved using a fairly standard method in hydrodynamics, the H_{-1} method, with technical modifications to deal with the lousy ergodicity properties of the model, and the singular behavior of the Stefan problem at creation points. The H_{-1} method was developed in [18] and [5] to study the metastability and nonequilibrium properties of conservative Ginzburg–Landau models. It is probably the easiest method in hydrodynamic scaling limits; however, it is restricted to gradient systems. In particular, it completely avoids the use of entropy arguments and the two-block estimate, which in the internal DLA models seem to be hopeless. In fact, the only type of ergodicity result needed as input is the characterization of the translation-invariant, invariant measures for the process in infinite volume which is proved using coupling methods of [1, 13]. This unfortunately forces us to assume the attractiveness (monotonicity) of the dynamics. We should also remark that the H_{-1} method as presented in this paper works for internal DLA if the nearest-neighbor walk is replaced by any symmetric finite-range random walk. Then one obtains ellipses as limiting shapes. Although one expects the same phenomena for mean-zero asymmetric models, they will require a different approach. The free random walk, or zero-range dynamics of the live particles can also be replaced by symmetric simple exclusions. Alternatively, the number of particles α frozen at a site can be replaced by a stationary random field. The method yields comparable results in all these cases.

1. Notation and results. We now describe the *generalized internal DLA* model. Particles are moving on the lattice $\varepsilon\mathbb{Z}^d$. We let η_x denote the number of particles at $x \in \varepsilon\mathbb{Z}^d$. The rate of jumping of particles from x to nearest-neighbor site y is $\varepsilon^{-2}a(\eta_x)$, where $a(n) = g(n-\alpha)$ for $n > \alpha$ and $a(n) = 0$ for $n \leq \alpha$ corresponding to the freezing of the first α particles. $g(n)$ are the rates of an attractive zero-range process and are assumed to satisfy, for some $C < \infty$,

$$C^{-1} \leq g(n+1) - g(n) \leq C$$

for all n. We will also assume that $g(n)$ is approximately linear, that is, $|g(n) - Kn| \leq C$ for all $n = 0, 1, \dots,$ for some finite C and K. Of course, by rescaling time we can and will assume $K = 1$ so the assumption is that, for all $n = 0, 1, \dots,$ for some finite C,

$$|g(n) - n| \leq C.$$

Particles are created at a finite number of sites $x_i = \lfloor \mathbf{x}_i \rfloor$ at rates c_i, $i = 1, \dots, n$. We use $\mathbf{x}$ to denote locations on $\mathbb{R}^d$, x to denote locations on $\varepsilon\mathbb{Z}^d$ and $\lfloor \mathbf{x} \rfloor$ to denote the closest point of $\varepsilon\mathbb{Z}$ to $\mathbf{x}$. The generator of the process is

$$L = \varepsilon^{-2} L_0 + \varepsilon^{-d} L_c,$$

where

$$L_0 f(\eta) = \sum_{x,e} a(\eta_x)\big(f(\eta^{x,x+e}) - f(\eta)\big), \tag{1.1}$$

where e denote the basis vectors of length ε in our lattice and $\eta_z^{x,y} = \eta_y + 1$ if $z = y$, $\eta_x - 1$ if $z = x$ and η_z otherwise, and

$$L_c f(\eta) = \sum_{i=1}^{n} c_i \big(f(\eta^{x_i}) - f(\eta)\big),$$

where $\eta_z^x = \eta_z + 1$ if $z = x$ and $\eta_z^x = \eta_z$ otherwise. The corresponding zero-range process, with generator

$$L_1 f(\zeta) = \sum_{x,e} g(\zeta_x)\big(f(\zeta^{x,x+e}) - f(\zeta)\big), \tag{1.2}$$

has as invariant measures the product measures with marginal probability that $\zeta_x = n$ given by

$$Z^{-1}(\gamma)\gamma^n / [g(n) \cdots g(1)]$$

for each x. The factor $Z(\gamma)$ is the normalization to make it into a probability measure. The parameter γ is related to the density ρ by the formula

$$\rho = \gamma Z'(\gamma)/Z(\gamma).$$

Define

$$\lambda(u) = \begin{cases} \gamma(u - \alpha), & \text{if } u \geq \alpha, \\ 0, & \text{if } 0 \leq u < \alpha. \end{cases}$$

Note that from the assumptions on $g(n)$ we have, for some finite C,

$$C^{-1}|u_1 - u_2| \le |\lambda(u_1) - \lambda(u_2)| \le C|u_1 - u_2|, \qquad u_1, u_2 \in [\alpha, \infty). \tag{1.3}$$

The upper bound actually holds trivially for all u_1, u_2. Note also that the assumptions imply that $|\gamma(\rho) - \rho| \le C$ for all ρ and therefore

$$|\lambda(u) - u| \le C \tag{1.4}$$

for all u.

The Stefan problem corresponding to our microscopic system is to find functions $\rho(\mathbf{x}, t)$ and $s(\mathbf{x})$ on $\mathbb{R}^d$ satisfying

$$\begin{cases} \dfrac{\partial \rho}{\partial t} = \Delta\gamma(\rho) + \displaystyle\sum_{i=1}^{n} c_i \delta_{\mathbf{x}_i}, & \text{in the region } \{\rho > 0\}, \\ \rho = 0, & \text{on } s(\mathbf{x}) \ge t, \\ \nabla_0 \rho \cdot \nabla s = -\alpha/\gamma'(0), & \text{on } s(\mathbf{x}) = t, \end{cases} \tag{1.5}$$

where $\nabla_0 \rho$ denotes the gradient taken from the inside of the region $s(\mathbf{x}) < t$. The initial condition is $\rho(\mathbf{x}, t) = 0$. Note the last equation can be rewritten $\nabla_0 \gamma(\rho) \cdot \nabla s = -\alpha$. The set $s(\mathbf{x}) \le t$ is the occupied set at time t, and the density $\rho(\mathbf{x}, t)$ evolves according to the nonlinear heat equation with Dirichlet boundary conditions on the occupied set, whose boundary in addition moves in the outward normal direction at a rate proportional to $\nabla\gamma(\rho)$. Physically, ρ corresponds to the density of live particles in our model and $\{(\mathbf{x}, t)\colon\ s(\mathbf{x}) \le t\}$ to the occupied region.

We make the following entropy transformation to obtain a weak formulation of the problem. Let ρ be a classical solution. Let φ be a smooth test function with compact support in $[0, T) \times \mathbb{R}^d$. Multiplying by φ and using Green's identity,

$$\int_0^T \int \left[\rho \frac{\partial \varphi}{\partial t} + \gamma(\rho)\Delta\varphi\right] d\mathbf{x}\, dt + \int_0^T \int_{s(x)=t} \varphi \nabla\gamma(\rho) \cdot \frac{\nabla s}{|\nabla s|}\, dS\, dt$$
$$+ \int_0^T \sum_{i=1}^{n} c_i \varphi(t, \mathbf{x}_i)\, dt = 0.$$

Using $\nabla\gamma(\rho) \cdot \nabla s = -\alpha$ on $s(\mathbf{x}) = t$, the middle term gives

$$-\alpha \int_0^T \int_{s(\mathbf{x})=t} \varphi \frac{dS}{|\nabla s|}\, dt = -\alpha \int \varphi(s(\mathbf{x}), \mathbf{x})\, d\mathbf{x} = \int_0^T \int \alpha 1_{\rho>0} \frac{\partial \varphi}{\partial t}\, d\mathbf{x}\, dt.$$

Let

$$\sigma(\rho) = \begin{cases} \rho + \alpha, & \text{if } \rho > 0, \\ 0, & \text{if } \rho = 0. \end{cases}$$

We conclude that

$$\int_0^T \int \left[\gamma(\rho)\Delta\varphi + \sigma(\rho)\frac{\partial \varphi}{\partial t}\right] d\mathbf{x}\, dt + \int_0^T \sum_{i=1}^{n} c_i \varphi(t, \mathbf{x}_i)\, dt = 0, \tag{1.6}$$

for all smooth test functions φ with compact support in $[0, T) \times \mathbb{R}^d$. Now let ρ satisfy (1.6) and define

$$u = \sigma(\rho).$$

Note that $\gamma(\rho) = \lambda(u)$. Thus u is a weak solution of

$$\frac{\partial u}{\partial t} = \Delta\lambda(u) + \sum_{i=1}^{n} c_i \delta_{\mathbf{x}_i} \tag{1.7}$$

in the sense that, for each smooth test function φ with compact support in $[0, T) \times \mathbb{R}^d$,

$$\int_0^T \int \left[\lambda(u)\Delta\varphi + u\frac{\partial\varphi}{\partial t}\right] d\mathbf{x}\, dt + \int_0^T \sum_{i=1}^{n} c_i \varphi(t, \mathbf{x}_i)\, dt = 0. \tag{1.8}$$

We will call such a ρ a weak solution of the Stefan problem. We will prove below an existence and uniqueness theorem for such weak solutions. Note that even in the simple case of two creation sites, at some time the clusters will meet and the weak solution will fail to satisfy (1.5) in a classical sense. We will not pursue the issue here of in what sense weak solutions satisfy (1.5) but refer the reader to the extensive literature (see, e.g., [16]).

The main results of this paper are the hydrodynamic limit for the particle density and the consequent shape theorems:

THEOREM 1.1. *For each $t > 0$, as $\varepsilon \to 0$, the empirical density field $\varepsilon^d \sum_{x \in \varepsilon\mathbb{Z}^d} \eta_x^\varepsilon(t)\delta_x$ of the generalized internal DLA process converges weakly in probability to the unique weak solution $u(t)$ of the Stefan problem* (1.5).

To state the shape theorems, let us introduce some notation. Let A_t^ε be the occupied set in the generalized internal DLA process with creation at rate ε^{-d} at the origin. Note that A_t^1 is in the occupied set when creation and diffusion occur at the same rate, hence it corresponds to the occupied set for the model in [3]. In all dimensions we denote by $B(\mathbf{x}, r)$ the Euclidean ball of radius r centered at $\mathbf{x}$.

DEFINITION. The distance $\text{dist}(A, \mathscr{A})$ between a set $A \subset \varepsilon\mathbb{Z}^d$ and a Borel set $\mathscr{A} \subset \mathbb{R}^d$ is defined as follows. Let $\overline{A} = \cup_{x \in \varepsilon\mathbb{Z}^d} \prod_{i=1}^d [x, x + \varepsilon e_i)$ be the natural embedding of A in $\mathbb{R}^d$, where e_i are the unit vectors in the positive coordinate directions. We define

$$\text{dist}(A, \mathscr{A}) = |\overline{A} \triangle \mathscr{A}|,$$

where $|\cdot|$ is Lebesgue measure and $\mathscr{A} \triangle \mathscr{B}$ denotes the symmetric difference $\mathscr{A} \cup \mathscr{B} - \mathscr{A} \cap \mathscr{B}$.

THEOREM 1.2. (i) *In any dimension d and for each $t > 0$,*

$$\text{dist}\big(A_t^\varepsilon, \mathscr{A}_t\big) \to 0$$

as $\varepsilon \to 0$ *in probability, where* $\mathscr{A}_t = \{\mathbf{x}\colon s(\mathbf{x}) \le t\}$ *is the occupied set in the Stefan problem* (1.5).

In the following results about the occupied set of Stefan problem (1.5), *we will assume that the creation of particles is at the origin, at rate* 1, *that is,* $n = 1$ *and* $\mathbf{x}_1 = 0$ *in* (1.5).

(ii) *Assume* $d = 1$. *There exist constants* K_1 *and* K_2 *such that, for sufficiently large* t,

$$B\big(0, K_1\sqrt{2t\log t}\big) \subset \mathscr{A}_t \subset B\big(0, K_2\sqrt{2t\log t}\big). \tag{1.9}$$

If $g(n) = n$ (*the case of independent random walks*), *then, for arbitrary* $\delta > 0$ *and* $K_1 = 1 - \delta$ *and* $K_2 = 1 + \delta$, (1.9) *holds for sufficiently large* t, *and it continues to hold with probability* 1 *if* $\mathscr{A}_t$ *is replaced with* A_t^1.

(iii) *Assume* $d = 2$. *There exists a constant* K *such that*

$$\mathscr{A}_t = B(0, K\sqrt{t})$$

and

$$\operatorname{dist}\big(A_t^1/\sqrt{t}, B(0, K)\big) \to 0$$

in probability, as $t \to \infty$. *In the case* $g(n) = n$, K *is the solution of*

$$\exp(-K^2/4) = \pi\alpha K^2.$$

(iv) *Assume* $d \ge 3$. *There exist constants* K_1 *and* K_2 *such that, for sufficiently large* t,

$$B\big(0, K_1 t^{1/d}\big) \subset \mathscr{A}_t \subset B\big(0, K_2 t^{1/d}\big). \tag{1.10}$$

If $g(n) = n$, *then, for arbitrary* $\delta > 0$ *and* $K_1 = (1 - \delta)a_d$ *and* $K_2 = (1 + \delta)a_d$, (1.10) *holds for sufficiently large* t, *and it continues to hold with probability* 1 *if* $\mathscr{A}_t$ *is replaced with* A_t^1. *Here,* $a_d = (d\Gamma(d/2)/2)^{1/d}/\sqrt{\pi}$, *where* Γ *is the gamma function.*

When $g(n) = n$, (1.10) for A_t^1 was proved already in [3]. (The random walks in this paper have different rates, but the argument applies unchanged.) In the case of zero–range dynamics, strengthening (1.9) or (1.10) to make K_1 and K_2 arbitrarily close remains an open problem, and so does proving (1.9) or (1.10) for A_t^1 instead of $\mathscr{A}_t$.

2. Stefan problem. In this section we prove some preliminary results about the Stefan problem which we will need. Although these results use fairly standard methods, we were not able to find any references in the vast literature on the Stefan problem (see references in [16]) dealing with the seemingly natural situation of delta function sources. The Stefan problem that arises from the internal DLA models that we are considering are nonstandard in two ways: the delta function sources and the nonlinearity of the heat equation inside the occupied region. These break the variational structure and this is why there is some work to do.

Let $P_x^\varepsilon(t)$ be the number of particles created at x up to time t in the internal DLA process. Of course, $P_x^\varepsilon(t)$ is just a Poisson process with rate $\varepsilon^{-d}c_i$ if $x = x_i$ and vanishes otherwise. Let Δ_ε be the lattice Laplacian on $\varepsilon\mathbb{Z}^d$,

$$[\Delta_\varepsilon\phi]_x = \varepsilon^{-2}\sum_{|e|=\varepsilon}[\phi_{x+e} - \phi_x].$$

We consider the following lattice approximation to the Stefan problem,

$$\frac{\partial u^\varepsilon}{\partial t} = \Delta_\varepsilon \lambda(u^\varepsilon) + dP^\varepsilon \tag{2.1}$$

on $\varepsilon\mathbb{Z}^d$, with $u^\varepsilon(0) = 0$. This system is a well-defined realization by realization of the Poisson process. We can and will identify u^ε with the function on $\mathbb{R}^d$ whose value on the box of side length ε centered at $x \in \varepsilon\mathbb{Z}^d$ is u_x^ε.

Let q be the solution of

$$\frac{\partial q}{\partial t} = \Delta q + \sum_{i=1}^n c_i \delta_{\mathbf{x}_i}, \tag{2.2}$$

with $q(0) = 0$. Note that (2.2) can be solved explicitly and the solution is $q(\mathbf{x}, t) = \sum_{i=1}^n \int_0^t p(\mathbf{x} - \mathbf{x}_i, t - s)c_i\, ds$, where $p(\mathbf{x}, t)$ is the heat kernel on $\mathbb{R}^d$.

The main result of this section is

THEOREM 2.1. *The solutions u^ε of the lattice version of the Stefan problem* (2.1) *converge weakly to the unique weak solution u of the Stefan problem* (1.7) *with $u(0) = 0$ satisfying*

$$\int_0^T \|u(t) - q(t)\|^2_{L^2(\mathbb{R}^d)}\, dt < \infty. \tag{2.3}$$

Furthermore, the convergence is strong away from the creation points $\mathbf{x}_1, \dots, \mathbf{x}_n$.

The theorem will be proved in Lemmas 2.2–2.8. The reason it is not quite straightforward is that the nice space in which to look for the limit equation to hold is H_{-1} and we have to face the fact that the Dirac delta functions are simply not there if $d \geq 2$. The same problem comes up when we try to prove uniqueness. The idea behind the uniqueness result is that, if u and v are two solutions of (1.7), the H_{-1} norm contracts at rate $\int (u-v)(\lambda(u)-\lambda(v))$, which is positive because λ is nondecreasing. But in dimensions $d \geq 4$, the solutions to our Stefan problem are not even L^2 functions due to the Green's-function–like singularities at the creation points. So we have to identify and subtract the singularities. In fact, the singularities behave in the same way as for the linear heat equation with delta function sources, which can be solved analytically (this is why we made the assumption that the jump rates are not far from linear). In dimensions 1 and 2, the solution is always in L^2 and grows, at rate $t^{1/2}$ in $d = 1$ and logarithmically in $d = 2$. In $d \geq 3$, one can check easily in the linear case $\gamma(u) = u$ that the solution grows to the solution of $\Delta\lambda(u_\infty) = -\sum_{i=1}^n c_i\delta_{\mathbf{x}_i}$, and in the nonlinear case one expects the same behavior. Let us make this more precise. We compare the solution of (2.1) with the "free" problem on $\varepsilon\mathbb{Z}^d$,

$$\frac{\partial q^\varepsilon}{\partial t} = \Delta_\varepsilon q^\varepsilon + dP^\varepsilon, \qquad q^\varepsilon(0) = 0. \tag{2.4}$$

The solution $q^\varepsilon_x(t)$ is given explicitly by $\sum_{i=1}^n \int_0^t p^\varepsilon_{x-x_i}(t-s)\,dP^\varepsilon_{x_i}(s)$, where $p^\varepsilon_x(t)$ is the heat kernel on $\varepsilon\mathbb{Z}^d$. For any $1 \le p < d/(d-2)$ if $d \ge 3$, any $1 \le p < \infty$ if $d = 2$ and any $1 \le p \le \infty$ if $d = 1$, and any $T \ge 0$, there exists a finite constant $C_p(T)$ so that, for all $\varepsilon > 0$ and $0 \le t \le T$,

$$E\left[\varepsilon^d \sum_x \left[q^\varepsilon_x(t)\right]^p\right] \le C_p(T). \tag{2.5}$$

It is also clear that, for any $\delta > 0$ and any $T > 0$, there exists a finite constant $C(\delta, T)$ such that, for all $\varepsilon > 0$ and $0 \le t \le T$,

$$E\left[\varepsilon^d \sum_{|x-x_i|\ge\delta} \left[q^\varepsilon_x(t)\right]^2\right] \le C(\delta, T). \tag{2.6}$$

Note that $\lim_{\delta\to 0} C(\delta, T) < \infty$ only in $d \le 3$. These computations are a little bit standard so we only sketch the idea. By the local limit theorem, $\lim_{\varepsilon\to 0,\, x\to\mathbf{x}} \varepsilon^{-d} p^\varepsilon_x(t) = (4\pi t)^{-d/2}\exp\{-\mathbf{x}^2/4t\}$, and from this and the well-known error estimates for the convergence, or more directly from discrete versions of the Nash inequality (see [6]), one can show that $\varepsilon^d \sum_x |\varepsilon^{-d} p^\varepsilon_x(t)|^p \le Ct^{-d(p-1)/2}$. Now one can use Jensen's inequality to take out the time integral against the Poisson process and obtain the desired estimates.

Let $\|\cdot\|_{-1}$ denote the H_{-1} norm on $f\colon \varepsilon\mathbb{Z}^d \to \mathbb{R}$ defined by

$$\|f\|^2_{-1} = \sup_\phi \left\{2\varepsilon^d \sum_x f_x\phi_x - \frac{\varepsilon^{d-2}}{2} \sum_{x,\,|e|=\varepsilon} \left[\phi_{x+e} - \phi_x\right]^2\right\}. \tag{2.7}$$

The supremum can be taken over functions ϕ with finite support. Note that the norm can only be finite if $\sum_x f_x = 0$. The H_{-1} norm can be adapted to remove this restriction, but we will not since all functions we deal with will automatically sum to 0. Let g_ε be the kernel of $-\Delta_\varepsilon^{-1}$, that is, at least for functions f with finite support which sum to zero,

$$\left[-\Delta_\varepsilon^{-1} f\right]_x = \varepsilon^d \sum_y g_\varepsilon(x-y) f_y; \tag{2.8}$$

then we also have

$$\|f\|^2_{-1} = \varepsilon^{2d} \sum_{x,\,y} g_\varepsilon(y-x) f_y f_x. \tag{2.9}$$

Equations (2.7) and (2.9) are equivalent definitions of the H_{-1} norm. If one is finite, then the other is as well and they are equal. In $d \ge 3$, g_ε is the *Green's function*, $g_\varepsilon(x) = \sum_{n=0}^\infty p_n(0, x)$, where p_n are the n step transition probabilities of a simple random walk on $\varepsilon\mathbb{Z}^d$. In $d \le 2$, it is the *potential kernel*, $g_\varepsilon(x) = \lim_{N\to\infty} \sum_{n=0}^N p_n(0, x) - p_n(x, x)$.

LEMMA 2.2. *Let u^ε be the solution of* (2.1) *and let q^ε be the solution of* (2.4). *Then for each $T > 0$, there exists a finite sure constant C independent of ε such that*

$$\int_0^T \|u^\varepsilon(t) - q^\varepsilon(t)\|_{\ell^2(\varepsilon\mathbb{Z}^d)}\, dt \le C. \tag{2.10}$$

PROOF. We have

$$\|u^\varepsilon(t) - q^\varepsilon(t)\|_{-1}^2 = -2\int_0^t \varepsilon^d \sum_x \big(u_x^\varepsilon - q_x^\varepsilon\big)\big(\lambda\big(u_x^\varepsilon\big) - q_x^\varepsilon\big)\, ds.$$

We use the bounds $|\lambda(u) - u| \le C$ and $\varepsilon^d \sum_x u_x^\varepsilon$, $\varepsilon^d \sum_x q_x^\varepsilon \le C$ to bound this by

$$-2\int_0^t \varepsilon^d \sum_x \big(u_x^\varepsilon - q_x^\varepsilon\big)^2\, ds + C.$$

We conclude that the absolute value of the first term is bounded by a constant independent of ε. $\square$

REMARK. In dimensions $d \le 3$, both terms in (2.10) are in ℓ^2 uniformly in ε so the lemma is easier. This is proved as follows. If $g_x = \int_0^t p_x^\varepsilon(t-s)\, ds$, then by Schwarz's inequality, $g_x^2 \le \int_0^t [p_x^\varepsilon(s)]^2 s^{-c}\, ds \int_0^t s^c\, ds$. Since $\|p^\varepsilon(s)\|_{\ell^2} \le Cs^{-d/2}$, we have that $\|g\|_{\ell^2}$ is finite if we can choose $-1 < c < (2-d)/2$, that is, if $d \le 3$. For (2.4), one has the integral against a Poisson process instead of Lebesgue measure, but the same result holds with probability 1. For (2.1), one has the added complication of the nonlinearity. In this case, the corresponding "free" problem is $\partial q^\varepsilon/\partial t = \Delta_\varepsilon \gamma(q^\varepsilon) + dP^\varepsilon$. We can write $\Delta_\varepsilon\gamma(q^\varepsilon) = \sum_\varepsilon a_{x,\,x+e}[u_{x+e}^\varepsilon - u_x^\varepsilon]$, where $a_{x,\,x+e} = [\gamma(q_{x+e}^\varepsilon) - \gamma(\rho_x^\varepsilon)]/(q_{x+e}^\varepsilon - q_x^\varepsilon)$ are bounded above and below. So $\partial q^\varepsilon/\partial t = \Delta_\varepsilon\gamma(q^\varepsilon)$ is the forward equation for a reversible random walk and therefore one knows that the transition probabilities are bounded above and below by Gaussians, by standard parabolic estimates for divergence form diffusion equations [6]. Therefore, the above argument shows that the solution of the inhomogeneous equation $\partial q^\varepsilon/\partial t = \Delta_\varepsilon\gamma(q^\varepsilon) + dP^\varepsilon$ is square integrable in dimensions 3 and lower. Finally, the same estimate holds for solutions of (2.1) by comparison with this case, because $u^\varepsilon \le q^\varepsilon + \alpha$.

LEMMA 2.3. *Let u^ε be the solution of* (2.1). *Then for each $\delta > 0$, there exists a constant $C(\delta) < \infty$ depending only on δ and u_0 such that*

$$E\left[\int_0^T \varepsilon^d \sum_{|x-x_i|\ge\delta} \sum_{|e|=\varepsilon} \varepsilon^{-2}\big\{\lambda\big(u_{x+e}^\varepsilon\big) - \lambda\big(u_x^\varepsilon\big)\big\}^2\, dt\right] \le C(\delta). \tag{2.11}$$

REMARK. The idea of the lemma is that if one has a solution of $\partial u/\partial t = \Delta\lambda(u)$, then the rate of contraction of the L^2 norm is $2\int \lambda'(u)|\nabla u|^2\, dx$, so the time integral of this term can be controlled.

PROOF. Let $J(\mathbf{x})$ be a smooth function vanishing in a neighborhood of the $\mathbf{x}_i$. We have

$$\varepsilon^d \sum_x \Big[u_x^\varepsilon(t)\Big]^2 J(x) = 2\int_0^t \varepsilon^d \sum_x u^\varepsilon \Delta_\varepsilon \lambda(u^\varepsilon) J\, ds.$$

After a summation by parts, the last term becomes

$$-\int_0^t \varepsilon^{d-2} \sum_{x,\,e} \nabla_{\varepsilon,\,e} u^\varepsilon \nabla_{\varepsilon,\,e} \lambda(u^\varepsilon) J\, ds - \int_0^t \varepsilon^{d-2} \sum_{x,\,e} u_{x+e}^\varepsilon \nabla_{\varepsilon,\,e} \lambda(u^\varepsilon) \nabla_{\varepsilon,\,e} J\, ds,$$

where $\nabla_{\varepsilon,e} f_x = f_{x+e} - f_x$. Applying Schwarz's inequality and using the assumptions on λ, we see that, for some $0 < c$ and $C < \infty$, this is bounded above by

$$-c\int_0^t \varepsilon^{d-2} \sum_{x,\,e} \big(\nabla_{\varepsilon,\,e} \lambda(u^\varepsilon)\big)^2 J\, ds + C\int_0^t \varepsilon^{d-2} \sum_{x,\,e} \big[u_{x+e}^\varepsilon\big]^2 \big(\nabla_{\varepsilon,\,e} J\big)^2 J^{-1}\, ds.$$

Now we choose a smooth J which is unity on $|\mathbf{x} - \mathbf{x}_i| \geq \delta$, vanishes on a neighborhood of the $\mathbf{x}_i$ and has $\varepsilon^{-2}(J(\mathbf{x}+e) - J(\mathbf{x}))^2 J^{-1}(\mathbf{x})$ uniformly bounded in ε. (Note this can be done since $f(r) = r^2$ satisfies $|f'|^2 \leq Cf$.) By the previous lemma and (2.6), u^ε have a uniform bound in L^2 in any region not containing the creation points $\mathbf{x}_i$. Hence the last term is bounded independent of ε. We conclude that the first term is also bounded independent of ε, which is what was to be proved. □

For fixed $\delta > 0$, define for random functions f on $\varepsilon\mathbb{Z}^d$, $\|f\|_B^2 = \sup E[\varepsilon^d \times \sum_x f_x \phi_x]$, where the supremum is over ϕ on $\varepsilon\mathbb{Z}^d$ with $\varepsilon^{d-2}\sum_{x,\,|e|=1} [\phi_{x+e} - \phi_x]^2 \leq 1$, $\|\phi\|_\infty \leq 1$, and $\phi = 0$ on $|x - x_i| \leq \delta$. Let B denote the corresponding Banach space. A comment on spaces is in order. Many of the spaces used in this paper are defined for functions on $\varepsilon\mathbb{Z}^d$ and may at first seem to depend rather strongly on ε. However, these are all really standard Sobolev spaces of functions on $\mathbb{R}^d$. We always use the natural map of a function f on $\varepsilon\mathbb{Z}^d$ to a function $\tilde{f}$ on $\mathbb{R}^d$ obtained by setting the latter equal to the former on a box of side length ε about the lattice point. To keep the connections to the particle systems concrete, we have decided to always write explicitly the Sobolev norms in terms of the original lattice functions.

LEMMA 2.4. *Let $g \leq \lambda$ be any smooth function with bounded first and second derivatives, and let u^ε be the solutions of the lattice version of the Stefan problem* (2.1). *Then $g(u^\varepsilon)$ are precompact as elements of the Banach space $L^2([0,T]; B)$ with norm $\int_0^T \|f(s)\|_B^2\, ds$.*

PROOF. By Rellich's theorem, it comes down to obtaining a uniform bound on

$$\int_0^T \left\| \frac{\partial}{\partial s} g\big(u^\varepsilon(s)\big) \right\|_B^2 ds.$$

(See [15] for a proof which applies in this context.) Take any ϕ with $\|\phi\|_\infty \le 1$, $\varepsilon^{d-2}\sum_{x,\,|e|=1}[\phi_{x+e}-\phi_x]^2 \le 1$ and $\phi = 0$ on $|x-x_i| \le \delta$. We have $(\partial/\partial t)\varepsilon^d \sum_x g(u_x^\varepsilon)\phi_x = \varepsilon^d \sum_x g'(u_x^\varepsilon)\Delta_\varepsilon \lambda(u_x^\varepsilon)\phi_x$. Summing by parts, this becomes

$$\varepsilon^{d-2}\sum_x g'(u_{x+e}^\varepsilon)\nabla_{\varepsilon,\,e}\lambda(u^\varepsilon)\nabla_{\varepsilon,\,e}\phi + \nabla_{\varepsilon,\,e}\lambda(u^\varepsilon)\nabla_{\varepsilon,\,e}g'(u^\varepsilon)\phi_x.$$

By Schwarz's inequality, the assumptions on ϕ and the bounds on g' and g'', we can control this by a constant plus a constant multiple of

$$\varepsilon^{d-2}\sum_{|x-x_i|\ge\delta,\,|e|=\varepsilon}\left[\nabla_{\varepsilon,\,e}\lambda(u^\varepsilon)\right]^2.$$

By the previous lemma, this establishes the precompactness. $\square$

From the L^2 or L^p bounds, and the fact that $\lambda(u) \le Cu$, it is easy to extract subsequences of u^ε and $\lambda(u^\varepsilon)$ converging weakly on $[0,T]\times\mathbb{R}^d$ to u and $\bar\lambda$, respectively. What is not straightforward is that the convergence is strong away from the creation points and that $\bar\lambda = \lambda(u)$. It is proved below.

LEMMA 2.5. *$\lambda(u^\varepsilon) \to \lambda(u)$ strongly.*

PROOF. Let ϕ be a smooth function which vanishes on any ball of radius $\delta/2$ around the creation points and is 1 on the complement of the union of balls of radius δ around the creation points. Let $p_a = p(\mathbf{x}, a)$ be the heat kernel on $\mathbb{R}^d$. Let $g \le \lambda$ be a smooth approximation to λ with two bounded derivatives. From the previous lemma, we can obtain a limit $\bar g$ in $L^2([0,T];B)$. In particular, for each fixed a,

$$\lim_{\varepsilon\to 0}\int_0^T\int |(\phi g(u^\varepsilon) - \phi\bar g) * p_a|^2\,d\mathbf{x}\,dt = 0.$$

It is easy to check that the inequality $\int |f * p_a - f|^2\,dx \le a\int|\nabla f|^2\,dx$ holds for any function for which it makes sense. Also, since $g \le \lambda$ is smooth with bounded derivative, Lemma (2.3) holds with $\lambda(u^\varepsilon)$ replaced by $g(u^\varepsilon)$ with a possible change of the constant $C(\delta)$ to $C(\delta, g)$. Together these give

$$\int_0^T\int|\phi g(u^\varepsilon) - \phi g(u^\varepsilon) * p_a|^2\,d\mathbf{x}\,dt \le Ca,$$

where $C < \infty$ again depends on g and δ but not on ε. When we pass to the weak limit, by lower semicontinuity of the norm we have $\int_0^T\int_{|x-x_i|\ge\delta}|\nabla\bar g|^2\,dx\,dt \le C(\delta, g)$ as well, so the above estimate holds with $g(u^\varepsilon)$ replaced by $\bar g$. We can therefore write

$$\begin{aligned}\int_0^T\int_{|x-x_i|\ge\delta}|g(u^\varepsilon)-\bar g|^2\,d\mathbf{x}\,dt \le{}& \int_0^T\int|\phi g(u^\varepsilon) - \phi g(u^\varepsilon) * p_a|^2\,d\mathbf{x}\,dt\\ &+\int_0^T\int|(\phi g(u^\varepsilon)-\phi\bar g) * p_a|^2\,d\mathbf{x}\,dt\\ &+\int_0^T\int|\phi\bar g * p_a - \phi\bar g|^2\,d\mathbf{x}\,dt.\end{aligned}$$

Each term vanishes in the limit as $\varepsilon \to 0$; then $a \to 0$. This proves the strong convergence of $g(u^\varepsilon)$ to $\bar{g}$ away from the creation points. Now we can choose $g(u)$ of the form described uniformly close to $\lambda(u)$ to conclude that $\lambda(u) \to \bar{\lambda}$ strongly away from the creation points as well.

Next we want to show that $\bar{\lambda} = \lambda(u)$. If $\lambda(u)$ had a Lipschitz inverse, this would follow directly from the strong convergence away from the creation points. However, $\lambda(u)$ is not even invertible, so it requires an argument. Recall that $\lambda(u) = \gamma(u - \alpha)$ for $u \geq \alpha$ and 0 for $u \leq \alpha$, where γ is a nice invertible function with $\gamma(0) = 0$. Suppose that f is a Lipschitz function with $f(0) = 0$ and $f(\lambda) \leq \gamma^{-1}(\lambda) + \alpha$. Because f is Lipschitz, we have that $f(\lambda(u))$ converges strongly to $f(\bar{\lambda})$. Because $f(\lambda) \leq \gamma^{-1}(\lambda) + \alpha$, we have that $f(\lambda(u^\varepsilon)) \leq u^\varepsilon$ for each $\varepsilon > 0$. Taking limits, we get $f(\bar{\lambda}) \leq u$. Taking the supremum of the left-hand side over such f, we get $\gamma^{-1}(\bar{\lambda}) + \alpha \leq u$ whenever $u > 0$. On the other hand, $u^\varepsilon \leq \gamma^{-1}(\lambda(u^\varepsilon)) + \alpha$ for each $\varepsilon > 0$, and taking limits gives $u \leq \gamma^{-1}(\bar{\lambda}) + \alpha$. The two inequalities together suffice to identify $\bar{\lambda} = \lambda(u)$. □

LEMMA 2.6. *u is a weak solution of the Stefan problem* (1.7) *satisfying* (2.3).

PROOF. Since we know now that $\lambda(u^\varepsilon) \to \lambda(u)$ strongly, we can simply take limits in the weak formulation of the microscopic Stefan problem to see that u is a weak solution of (1.7). From (2.10) we obtain

$$\int_0^T \|u^\varepsilon(t) - q^\varepsilon(t)\|^2_{\ell^2(\varepsilon\mathbb{Z}^d)}\, dt \leq C.$$

Since $u^\varepsilon - q^\varepsilon$ converge to $u - q$ weakly, (2.3) follows by lower semicontinuity. □

LEMMA 2.7. *There is at most one solution of the Stefan problem* (1.7) *satisfying* (2.3).

PROOF. Suppose u and v are two such solutions. From (2.3) we obtain

$$\int_0^T \|u(t) - v(t)\|^2_{L^2(\mathbb{R}^d)}\, dt < \infty.$$

By a well-known theorem [14], if $V \subset H \subset V'$ are separable Hilbert spaces with V and V' being in duality relative to the inner product of H and we have a curve $u(t)$: $0 \leq t \leq T$ satisfying $u \in L^2(0, T; V)$ and $u' \in L^2(0, T; V')$, then u is almost everywhere equal to a function continuous from $[0, T]$ into H and for $0 \leq t_1 \leq t_2 \leq T$, $\|u(t_2)\|_H^2 - \|u(t_1)\|_H^2 = 2\int_{t_1}^{t_2} \langle u(t), u'(t)\rangle\, dt$. Therefore,

$$\|u(T) - v(T)\|^2_{H_{-1}(\mathbb{R}^d)} \leq -2\int_0^T \int (u - v)(\lambda(u) - \lambda(v))\, d\mathbf{x}\, dt.$$

Since λ is nondecreasing, we conclude that $\|u(T) - v(T)\|_{H_{-1}(\mathbb{R}^d)} = 0$. This is true for every $T > 0$ and therefore $u \equiv v$. □

LEMMA 2.8. *$u^\varepsilon \to u$ strongly away from the creation points.*

PROOF. First note that for our solution, the Lebesgue measure of the *mushy region* $0 < u \le \alpha$ must be zero, because if u is a weak solution, then so is $u - u\mathbf{1}_{\{0<u\le\alpha\}}$ and then uniqueness tells us that $u\mathbf{1}_{\{0<u\le\alpha\}} \equiv 0$.

Now let $f_n(u) = u$ if $u \ge \alpha + 1/n$, $f_n(u) = 0$ if $u \le \alpha$ and f_n is linear in between. Each function f_n is Lipschitz continuous and they approximate the function $f(u) = u\mathbf{1}(u > \alpha)$. We write $u^\varepsilon = f_n(u^\varepsilon)+(u^\varepsilon - f_n(u^\varepsilon))$. By the strong convergence of $\lambda(u^\varepsilon)$ to $\lambda(u)$ away from creation points, we have that $f_n(u^\varepsilon)$ converges strongly to $f_n(u)$ as $\varepsilon \to 0$ there, too. Since u^ε converges weakly to u, the final term $u^\varepsilon - f_n(u^\varepsilon)$ converges weakly to $u - f_n(u)$. By the previous paragraph, this converges to 0 as $n \to \infty$. Choosing a diagonal sequence, we can therefore have $f_{n(\varepsilon)}(u^\varepsilon) \to u\mathbf{1}(u > \alpha)$ strongly and $u^\varepsilon - f_{n(\varepsilon)}(u^\varepsilon) \to 0$ weakly. However, all terms are nonnegative, and a nonnegative sequence converging weakly to zero, must in fact converge strongly. From this we infer the strong convergence of the sequence u^ε away from the creation points, and this completes the proof. $\square$

One of the nice consequences of having the lattice approximation to the Stefan problem is that it allows us to prove very easily comparison principles for solutions even though the creation rates are singular. For example, passing to the limit after an easy computation on the lattice level using the maximum principle, we obtain the following result which will be useful in Section 4.

LEMMA 2.9. *Let u be the weak solution of* (1.7) *with bounded initial condition u_0 and let v be a subsolution* (*respectively*, *supersolution*) *satisfying*

$$\frac{\partial v}{\partial t} - \Delta\lambda(v) \le \sum_{i=1}^{n} c_i \delta_{\mathbf{x}_i}, \quad v(0, \mathbf{x}) = v_0(\mathbf{x}) \qquad (resp.\ \ge). \tag{2.12}$$

If $v_0 \le u_0$, then $v \le u$(resp. $\ge$).

3. Hydrodynamic limit. Recall the definition (2.7), (2.9) of the H_{-1} norm $\|\cdot\|_{-1}$ on $\varepsilon\mathbb{Z}^d$. The main result of this section is

THEOREM 3.1. *Let η^ε be the configuration of the generalized internal DLA process and let u^ε be the solution of the discretized Stefan problem* (2.1) *with the same initial condition*. *Then*

$$\lim_{\varepsilon\to 0} E\big[\|\eta^\varepsilon(T) - u^\varepsilon(T)\|_{-1}^2\big] = 0.$$

Here and in the rest of the section, E denotes expectation with respect to the generalized internal DLA process $\eta^\varepsilon(\cdot)$. The theorem will be proved in Lemmas (3.2)–(3.11).

LEMMA 3.2. *Let η^ε be the configuration of the generalized internal DLA process and let u^ε be the solution of the discretized Stefan problem* (2.1) *with* *the same initial condition. Then*

$$\|\eta^\varepsilon(T) - u^\varepsilon(T)\|_{-1}^2 = 2\int_0^T \varepsilon^d \sum_x V(\eta_x^\varepsilon, u_x^\varepsilon)\, dt + M^\varepsilon(T), \tag{3.1}$$

where

$$V(n, u) = -(n-u)(a(n) - \lambda(u)) + a(n) \tag{3.2}$$

and $M^\varepsilon(t)$ is a martingale.

PROOF. This is an explicit computation. We compute

$$\begin{aligned} d[(\eta_x - u_x)(\eta_y - u_y)] = \Big\{&[\Delta_\varepsilon(a(\eta_x) - \lambda(u_x))](\eta_y - u_y) \\ &+ (\eta_x - u_x)[\Delta_\varepsilon(a(\eta_y) - \lambda(u_y))] \\ &- \varepsilon^{-2}\mathbf{1}(|x-y| = 1)(a(\eta_x) + a(\eta_y)) \\ &+ \varepsilon^{-2}\mathbf{1}(x = y)\sum_e (a(\eta_x) + a(\eta_{x+e}))\Big\}\, dt + dM_{xy}, \end{aligned}$$

where M_{xy} are martingales, and then use the fact that $[-\Delta_\varepsilon g^\varepsilon]_0 = \varepsilon^{-d}$. $\square$

Now consider the generalized IDLA process running in a finite box Λ with reflecting boundary conditions and no creation. The set of invariant measures will be denoted μ_Λ^β running over the parameter set $\beta \in B$. They are described as follows. For each $N \geq \alpha|\Lambda|$, there is exactly one invariant measure with N particles. Under that measure, each $\eta_x \geq \alpha$ and the configuration $\eta_x - \alpha$, $x \in \Lambda$ is distributed according to the canonical invariant measure for our reference zero-range process with $N - \alpha|\Lambda|$ particles. On the other hand, if $N < \alpha|\Lambda|$, then for each configuration of N particles in Λ with no more than α particles at each site, the Dirac mass at that configuration is invariant. So the parameter set B consists really of each fixed configuration with no more than α particles per site, together with each $N \geq \alpha|\Lambda|$.

LEMMA 3.3. *The marginal on Λ of any translation-invariant, invariant measure for L_0 on $\mathbb{Z}^d$ is a mixture of the μ_Λ^β.*

The proof of this lemma is given in Section 5. We now continue with the proof of Theorem 3.1

LEMMA 3.4. *Let V be as in* (3.2) *and let μ be a translation-invariant, invariant measure for the process. Then for any u,*

$$E_\mu[V(\eta_0, u)] \leq 0.$$

PROOF. Let Λ be a box around the origin and let μ_Λ be the marginal of μ on Λ. By the previous lemma, μ_Λ is a mixture of the extremal invariant measures μ_Λ^β on Λ. Let $f(n)$ be any function of whole numbers which vanishes if $n \leq \alpha$

and let μ_Λ^β be an extremal invariant measure on Λ with density v_β. Let ν_γ be an invariant measure for the zero-range dynamics, that is, a product measure with marginals $\nu_\gamma(\eta_x = n) = Z^{-1}(\gamma)\gamma^n[g(n)\cdots g(1)]^{-1}$. We have

$$E_{\mu_\Lambda^\beta}[f(\eta_0)] = \begin{cases} 0, & \text{if } v_\beta \leq \alpha, \\ E_{\nu_{\gamma(v_\beta-\alpha)}}[f(\eta_0-\alpha)] + O(|\Lambda|^{-1}) & \text{if } v_\beta > \alpha. \end{cases}$$

In the second case, the error $O(|\Lambda|^{-1})$ comes from changing the measure μ_Λ^β with fixed density to the product measure ([11, section 6] is a good reference for the equivalence of ensembles in this context). Note that both $a(n)$ and $na(n)$ vanish if $n \leq \alpha$. Furthermore, if $v > \alpha$,

$$E_{\nu_{\gamma(v-\alpha)}}[a(\eta_0)] = \lambda(v),$$
$$E_{\nu_{\gamma(v-\alpha)}}[\eta_0 a(\eta_0)] = (v+1)\lambda(v).$$

Letting $|\Lambda| \to \infty$, we have

$$E_\mu[V(\eta_0, u)] = -\int (v_\beta - u)(\lambda(v_\beta) - \lambda(u))\, d\Upsilon_\beta,$$

where Υ_β is a probability measure on the parameter set for the extremal invariant measures. This is nonpositive since λ is nondecreasing. Note the apparent lack of uniformity in the argument is easily resolved: except for small densities, the extremal invariant measures on Λ are parameterized by the density itself, and because of the stationarity of μ one has an easy cutoff of large densities uniform in Λ. $\square$

Consider the right-hand side of (3.1). If we take the expectation, the martingale disappears and because of the time averaging we will be evaluating $V(\eta_x, u)$ under an invariant measure. In fact, we can do some spatial averaging as well so that the resulting measure is also translation invariant. By Lemma 3.4, the result is nonpositive, giving the hydrodynamic limit. This argument is made precise in Lemma 3.10. However, there are several cutoffs needed first. Lemma 3.5 provides a basic estimate that will be used repeatedly. Lemma 3.6 cuts off the bad points near the creation sites where the density is unbounded. Lemma 3.7 cuts off large values of V, and Lemma 3.8 provides for the averaging in space and time in order to obtain a translation-invariant, invariant measure when we look at the limit of the distribution around a macroscopic site as we let $\varepsilon \to 0$.

LEMMA 3.5. (i) *Let* $1 \leq p \leq 2$ *if* $d \leq 3$ *and* $1 \leq p < d/(d-2)$ *if* $d \geq 4$. *For each* $T > 0$, *there is a constant* $C_p(T) < \infty$ *such that, for all* $\varepsilon > 0$,

$$E\left[\int_0^T \varepsilon^d \sum_x \left[\eta_x^\varepsilon(t)\right]^p dt\right] \leq C_p(T).$$

(ii) *There exists a function* $C(\delta) < \infty$ *for each* $\delta > 0$ *such that, for all* $\varepsilon > 0$,

$$E\left[\int_0^T \varepsilon^d \sum_{|x-x_i|\geq\delta} \left[\eta_x^\varepsilon(t)\right]^2 dt\right] \leq C(\delta).$$

PROOF. Follows from the similar bound on u and Lemma 3.1 which implies that the expectation of the L^2 norm of the difference of u and η can be controlled. □

LEMMA 3.6. *Let* V *be as in* (3.2). *Then*

$$E\left[\|\eta^\varepsilon(T) - u^\varepsilon(T)\|_{-1}^2\right] \leq 2\varepsilon^d \sum_{|x-x_i|\geq\delta} \int_0^T E\left[V\left(\eta_x^\varepsilon, u_x^\varepsilon\right)\right] dt + \Omega_1(\delta, \varepsilon, T), \tag{3.3}$$

where for each $T > 0$,

$$\limsup_{\delta\to 0} \limsup_{\varepsilon\to 0} \Omega_1(\delta, \varepsilon, T) = 0.$$

PROOF. Let $\Omega_1(\delta, \varepsilon, T) = CE[\int_0^T \varepsilon^d \sum_{\{x:\, |x-x_i|<\delta\}} \eta_x\, dt]$, where C is large enough that $a(n) \leq Cn$. Replacing $a(n)$ by $\lambda(n)$ in (3.2) gives something negative and the error can be bounded by a finite multiple of n. We can estimate Ω_1 by part 1 of Lemma 3.5 and Chebyshev's inequality. □

LEMMA 3.7. *Let* $\phi_\ell(x) = x$ *if* $x \leq \ell$ *and* $\phi_\ell(x) = \ell$ *if* $x \geq \ell$ *be the cutoff at level* ℓ. *Let*

$$V_\ell(n, x) = -\phi_\ell\big((n - x)(a(n) - \lambda(x))\big) + \phi_\ell(a(n)). \tag{3.4}$$

Then

$$\begin{aligned} &E\left[\|\eta^\varepsilon(T) - u^\varepsilon(T)\|_{-1}^2\right] \\ &\quad\leq 2\varepsilon^d \sum_{|x-x_i|\geq\delta} \int_0^T E\left[V_\ell\left(\eta_x^\varepsilon(t), u_x^\varepsilon(t)\right)\right] dt + \Omega_2(\ell, \varepsilon, \delta, T), \end{aligned} \tag{3.5}$$

where

$$\limsup_{\delta\to 0} \limsup_{\ell\to\infty} \limsup_{\varepsilon\to 0} \Omega_2(\ell, \varepsilon, \delta, T) = 0.$$

PROOF. Since one of the error terms from (3.3) is negative, all we have to show is that

$$\limsup_{\ell\to 0} \limsup_{\varepsilon\to 0} E\left[\int_0^T \varepsilon^d \sum_x |\phi_\ell(a(\eta_x)) - a(\eta_x)|\, dt\right] = 0,$$

which follows from Lemma 3.5 and Chebyshev's inequality. □

LEMMA 3.8. *Let V_ℓ be as in (3.4) and let u be the unique solution of the Stefan problem (1.5). For each $\sigma > 0$, let $B_\sigma(\mathbf{y}, t) = \{(x, s) \in \varepsilon\mathbb{Z}^d \times [0, T]: |x - \mathbf{y}| \le \sigma, |s - t| \le \sigma\}$. Then*

$$\begin{aligned} &E\big[\|\eta^\varepsilon(T) - u^\varepsilon(T)\|_{-1}^2\big] \\ &\quad \le 2\int_{|\mathbf{y}-x_i|\ge\delta}\int_0^T Av_{(x,s)\in B_\sigma(\mathbf{y},t)} E\big[V_\ell\big(\eta_x^\varepsilon(s), u(\mathbf{y}, t)\big)\big]\,dt\,d\mathbf{y} \qquad (3.6)\\ &\qquad + \Omega_3(\varepsilon, \sigma, \ell, \delta, T), \end{aligned}$$

where for each $T > 0$,

$$\limsup_{\delta\to 0}\limsup_{\sigma\to 0}\limsup_{\ell\to\infty}\limsup_{\varepsilon\to 0}\Omega_3(\varepsilon, \sigma, \ell, \delta, T) = 0.$$

Here Av denotes the average.

PROOF. After a summation by parts, the error term becomes

$$\Omega_3 = \int\int_0^T Av_{(x,s)\in B_\sigma(\mathbf{y},t)} E\big[V_\ell\big(\eta_x^\varepsilon(s), u(\mathbf{y}, t)\big) - V_\ell\big(\eta_x^\varepsilon(s), u_x^\varepsilon(s)\big)\big]\,dt\,d\mathbf{y}.$$

The $\mathbf{y}$ integral stays away from the creation sites $\mathbf{x}_i$ by $\delta - \sigma$. From the special form of V, and using the bounds $a(n) \le Cn$ and $\lambda' \le C$, and Schwarz's inequality, for each $\zeta > 0$,

$$\begin{aligned} &\big[V_\ell\big(\eta_x^\varepsilon, u(\mathbf{y})\big) - V_\ell\big(\eta_x^\varepsilon, u_x^\varepsilon\big)\big] \\ &\quad \le C\zeta^{-1}\big(\big(\eta_x^\varepsilon\big)^2 + \big(u_x^\varepsilon\big)^2 + u(\mathbf{y})^2\big) + \zeta\big(u(\mathbf{y}) - u_x^\varepsilon\big)^2. \end{aligned} \qquad (3.7)$$

Since $\delta > 0$ goes to zero last, we can use the uniform L^2 bounds on η^ε and u^ε away from the creation sites [(2.6) and Lemma 3.5] to show that the part of Ω_3 coming from the first term on the right-hand side of (3.7) is bounded by some $C'\zeta^{-1}$ uniformly in ε and σ for each $\delta > 0$. The second term vanishes in the limit as ε, then σ are sent to zero, by the strong convergence of the lattice approximations to the Stefan problem away from the creation points (Theorem 2.3). Finally, we can send $\zeta \to \infty$ to obtain the lemma. □

The following lemma provides us cheaply with the cutoff of large space that we will need.

LEMMA 3.9. *There are constants $C_1 < \infty$ and $C_2 > 0$ depending only on the walk rate $g(\cdot)$, the creation rates c_i and on T such that, for all $\varepsilon > 0$,*

$$E\bigg[\varepsilon^d\sum_{|x|\ge M}\eta_x(T)\bigg] \le C_1 M^{-1}\exp\{-C_2 M^2\}.$$

PROOF. Let X_t^i denote the positions at time t of all the particles with the convention that a particle is held at its creation site up to the time it is created. Clearly, the left-hand side is just the normalized number of such particles outside of $|x| \le M$ at time T. If we let $Y_t^i = X_t^i$ up to the time that X^i reaches its resting place, and then walk at rate ε^{-2} after that, then the left-hand side is clearly bounded by $\varepsilon^d E\big[\#i \mid \sup_{0\le t\le T} |Y^i_t| \ge M\big]$ which, since there are less than $C_3\varepsilon^{-d}$ such particles, is clearly bounded by

$$C_3 \sup_i P\bigg(\sup_{0\le t\le T} |Y^i_t| \ge M\bigg).$$

Let t_i be the creation time of particle i and note that after its creation it behaves as a symmetric random walk with jump rate $\varepsilon^{-2}a_t$, where a_t is bounded above and below away from zero and adapted to the σ-field of the entire process. Make a random time change $\tau_i(t)$ after time t_i so that, after its creation $Y^i_{\tau_i(t)}$ has jump rate ε^{-2}. Then the probability can be rewritten in terms of a standard random walk X_t running at rate ε^{-2} as $P\big(\sup_{0\le t\le \tau_i^{-1}(T)-t_i} |X_t| \ge M\big)$, which, because $\tau_i(t)$ is bounded below by $C_4 t$ for some $C_4 > 0$, is bounded above by

$$P\bigg(\sup_{0\le t\le C_5 T} |X_t| \ge M\bigg)$$

for some finite C_5. It is now standard to show that this is controlled by the term on the right-hand side of the lemma. □

From now on it will be convenient to think of our configuration η on $\mathbb{Z}^d$ instead of $\varepsilon\mathbb{Z}^d$ because we want to study the limiting measure. Note, however, that time is still sped up by a factor of ε^{-2}. Let $\mu^\varepsilon(t, d\eta)$ be the distribution of the η at time t and let

$$\bar{\mu}_{\varepsilon,\sigma}(\mathbf{y}, t, d\eta) = Av_{(x,s)\in B_\sigma(\mathbf{y},t)}\tau_x\mu^\varepsilon(s, d\eta), \tag{3.8}$$

where τ_x is the shift acting on measures: $\tau_x\mu(f(\eta)) = \mu(f(\tau_x\eta))$, where $[\tau_x\eta]_y = \eta_{x+y}$. Then we can write

$$Av_{(x,s)\in B_\sigma(\mathbf{y},t)} E\big[V_\ell(\eta_x(s), u(\mathbf{y}, t))\big] = E_{\bar{\mu}_{\varepsilon,\sigma}(\mathbf{y},t)}\big[V_\ell(\eta_0, u(\mathbf{y}, t))\big]. \tag{3.9}$$

Here we have used the notation $B_\sigma(\mathbf{y}, t)$ to be consistent with the statement of Lemma 3.8, but now sites x are on the unscaled lattice $\mathbb{Z}^d$ so the side length of the box is really $\sigma\varepsilon^{-1}$.

LEMMA 3.10. *For each $\mathbf{y} \in \mathbb{R}^d$ with $|\mathbf{y} - \mathbf{x}_i| \ge \delta$ and any $t \in [0, T]$ and $0 < \sigma < \delta$, the set of measures $\bar{\mu}_{\varepsilon,\sigma}(\mathbf{y}, t)$, $\varepsilon > 0$, defined in* (3.8) *is tight with respect to the weak topology on the set of configurations $\mathbb{N}^{\mathbb{Z}^d}$ and any weak limit is a translation-invariant*, *invariant measure of the internal DLA process* (*without creation*).

PROOF. Let c_x, $x \in \mathbb{Z}^d$, be any positive weights with $\sum c_x = 1$. The sets $\sum c_x\eta_x \le M$ are weakly compact and expand to fill the whole configuration space. By Chebyshev's inequality,

$$\bar{\mu}_{\varepsilon,\sigma}(\mathbf{y}, t)\bigg(\sum_x c_x\eta_x > M\bigg) \le M^{-1}E_{\bar{\mu}_{\varepsilon,\sigma}(\mathbf{y},t)}\bigg[\sum_x c_x\eta_x\bigg] \le CM^{-1}\sigma^{-d}$$

from the fact that the expected number of particles is less than some constant multiple of ε^{-d}. This establishes the tightness.

Since it is constructed as the limit of large averages of bounded measures, the translation invariance of the limit measure is immediate.

We now prove that such a limit measure μ must be invariant. Let f be any bounded local function and let L_0 be the generator of the internal DLA process without creation and without the speed up of time (1.1).

$$E_{\mu}[L_0 f(\eta)] = \lim_{\varepsilon\to 0} Av_{(x,\,s)\in B_{\sigma}(\mathbf{y},t)} E\big[L_0 f(\tau_x \eta_s)\big].$$

By the translation invariance of the dynamics,

$$Av_{|x-\mathbf{y}|\le\sigma\varepsilon^{-1}} E[L_0 \tau_x f] = E\big[L_0\big\{Av_{|x-\mathbf{y}|\le\sigma\varepsilon^{-1}}\tau_x f\big\}\big].$$

By the definition of the generator,

$$\begin{aligned}\frac{1}{2\sigma}\int_{t-\sigma}^{t+\sigma} & E\big[L_0\big\{Av_{|x-y|\le\sigma\varepsilon^{-1}}\tau_x f\big\}(s)\big]\,ds \\ &= \frac{\varepsilon^2}{2\sigma}E\big[Av_{|x-y|\le\sigma\varepsilon^{-1}}\tau_x f\big]\Big|_{t-\sigma}^{t+\sigma} \\ &\quad - \frac{\varepsilon^{2-d}}{2\sigma}\int_{t-\sigma}^{t+\sigma} E\big[L_c\big\{Av_{|x-y|\le\sigma\varepsilon^{-1}}\tau_x f\big\}\big]\,ds.\end{aligned}$$

Note, however, that the last term always vanishes for $\sigma < \delta$ and small enough ε since the support of $\tau_x f$ never contains any of the creation sites. Therefore, the right-hand side is bounded by $C\varepsilon^2$. Therefore, we have $E_{\mu}[L_0 f] = 0$ for any bounded local function and we conclude that μ is invariant. $\square$

LEMMA 3.11. $\limsup_{\varepsilon\to 0} E\big[\|\eta^{\varepsilon}(T) - u^{\varepsilon}(T)\|_{-1}^2\big] = 0.$

PROOF. By Lemma 3.8, we have that the left-hand side is bounded above by

$$\limsup_{\substack{\delta,\,\sigma\to 0\\ \delta>\sigma}}\limsup_{\ell\to\infty}\limsup_{\varepsilon\to 0}\int\int_0^T \mathbf{1}_{\{|\mathbf{y}-\mathbf{x}_i|\ge\delta\}} E_{\bar{\mu}_{\varepsilon,\,\sigma}(\mathbf{y},\,t)}\big[V_{\ell}(\eta_0, u(\mathbf{y},t))\big]\,dt\,d\mathbf{y}.$$

The first thing we want to do is pass the limit in ε through the integral. It is simple to check from the assumptions on $a(\cdot)$ that V_{ℓ} is the sum of a nonpositive term and a term which is bounded by a constant multiple of η_0. For the integral corresponding to the nonpositive term, commutation of the limit in ε and the integration in t and y is a direct application of Fatou's lemma. For the term which is bounded by a constant multiple of η_0, we break it into the integration over $|y| < M$ and over $|y| \ge M$. For fixed M, we can pass the ε limit through the first piece by the bounded convergence theorem, since V_{ℓ} is bounded for each fixed ℓ. Then we can use Lemma 3.9 to estimate away the second piece as $M\to\infty$. By the previous lemma, we therefore have that the left-hand side of the lemma is bounded above by

$$\limsup_{\substack{\delta,\,\sigma\to 0\\ \delta>\sigma}}\limsup_{\ell\to\infty}\int\int_0^T \mathbf{1}_{\{|\mathbf{y}-\mathbf{x}_i|\ge\delta\}} E_{\mu(\mathbf{y},\,t)}[V_{\ell}(\eta_0, u(\mathbf{y},t))]\,dt\,d\mathbf{y},$$

where for each $\mathbf{y}$ and t, $\mu(\mathbf{y}, t)$ is a translation-invariant, invariant measure for the dynamics without creation. Referring to (3.4), we can see that V_ℓ is the sum of two monotonic functions in ℓ. By the monotone convergence theorem, and by part 2 of Lemma 3.5, which bounds each term separately, we can take the limit in ℓ through the time and space integrations. The resulting integrand is the expectation of V with respect to a translation-invariant, invariant measure for the dynamics without creation, which vanishes by Lemma 3.4. This completes the proof of the hydrodynamic limit. $\square$

4. Shape theory. This section is devoted to convergence results for the occupied sets in the internal DLA and Stefan problems. We will denote by $A_t^\varepsilon \subset \varepsilon\mathbb{Z}^d$ the occupied set $\{x\colon \eta_x > 0\}$ and by $\mathscr{A}_t = \{\mathbf{x}\colon \rho(\mathbf{x}, t) > 0\}$ the occupied set of the Stefan problem (1.5). Recall that we defined the distance between a set on $\varepsilon\mathbb{Z}^d$ lattice and a set in Euclidean space to be the Lebesgue measure of the symmetric difference of the natural embedding of the first with the second.

THEOREM 4.1. *For every $t > 0$,*

$$\lim_{\varepsilon\to 0} \operatorname{dist}\big(A_t^\varepsilon, \mathscr{A}_t\big) = 0$$

in probability.

REMARK. What follows immediately from the hydrodynamic limit is that macroscopic densities on the occupied set A_t^ε and the occupied set in the Stefan problem $\mathscr{A}_t$ agree. On the other hand, convergence in dist means (1) that on the complement of a slightly larger set than the occupied set in the Stefan problem, the density of particles is going to zero and, (2) that on a slightly smaller set than the occupied set in the Stefan problem, the density of "holes" is going to zero. Note that the first statement follows immediately from the weak convergence, but the second does not because the density could be macroscopically correct while on the microscopic scale there could be clusters of free particles surrounded by seas of holes. The proof of Theorem 4.1 given below utilizes a bootstrapping argument to demonstrate that this does not happen. Actually, computer simulations leave no doubt that $A_t^1/\sqrt{t} \to \mathscr{A}_1$ a.s. in the even stronger Hausdorff metric.

LEMMA 4.2. *Let $\Lambda \subset \mathbb{Z}^d$ be a box of side length $L = a\varepsilon^{-1}$ centered at 0 and let Γ be a concentric box of side length $\Lambda/2$. Let $0 < \kappa_1 < 1$ and $0 < \kappa_2 < \infty$ be fixed. Let $H \subset \Lambda$ of cardinality $\kappa_1 L^d$. Let $N = \kappa_2 L^d$ particles start at $x_1, \dots, x_N \in \Gamma$ and perform zero-range dynamics with killing on the boundary. In other words, the generator of the process is $Lf(\eta) = \varepsilon^{-2}\sum_{x,e} g(\eta_x)(f(\eta^{x,x+e}) - f(\eta))$, where the sum is over $x \in \Lambda$ and $|e| = 1$ with the convention that if $x \in \Lambda$ and $x + e \notin \Lambda$, then $\eta_y^{x,x+e} = \eta_y$ unless $y = x$ in which case $\eta_x^{x,x+e} = \eta_x - 1$. Assume that for some $c < \infty$, $g(n) \le cn$. Let $p_{x_1,\dots,x_N,H}(\varepsilon)$ denote the probability that none of the particles are in H at time a^2. Let $p(\varepsilon) = \sup_{x_1,\dots,x_N,H} p_{x_1,\dots,x_N,H}(\varepsilon)$. Then $\limsup_{\varepsilon\to 0} \varepsilon^d \log p(\varepsilon) < 0$.*

PROOF OF LEMMA 4.2. Suppose first of all that the particles move independently. In other words, $g(n) = cn$ for some c. It is not hard to show that there exists $\gamma > 0$ so that the probability for a particle starting in Γ and killed on the boundary of Λ to be in H at time a^2 is at least γ, uniformly over all such initial points x and target sets H. We could call such an event a success. By standard large deviation results for any $r < 1$, the probability p_N to have fewer than $rN\gamma$ successes satisfies $\lim_{N\to\infty} N^{-1} \log p_N < 0$, and the probability we seek is certainly smaller.

Now we do the same computation in a different way again assuming that $g(n) = cn$. Define the empirical density field $\rho^\varepsilon(t, x)$ to be the number of particles at x at time t on the tile of side length ε containing εx. Let P_ε be the distribution of the field $\rho^\varepsilon(t)$, $0 \le t < \infty$. The event whose probability we want to compute is a subset of the event A that, given initial mass concentrated in $[-a/4, a/4]^d$, the empirical density field vanishes at time a^2 on some subset $H \subset [-a/2, a/2]^d$ of Lebesgue measure $\kappa_1 a^d > 0$. Fix H and note that, for each ε, $\sup_{x_1, \dots, x_N} p_{x_1, \dots, x_N, H}(\varepsilon)$ is achieved at some N points in Γ, so let us then start the system with the N particles at those points. The corresponding set of initial measures $\rho^\varepsilon(x, 0)\, dx$ will be tight. Suppose that for some subsequence ε_k, we have $\rho^{\varepsilon_k}(0, x)\, dx \Rightarrow \mu_0$ for some measure μ_0. Since the particles move independently, it is straightforward to show that then $\rho^{\varepsilon_k}(t, x)dx \Rightarrow \rho(t, x)dx$, where ρ is the solution of the linear heat equation $\partial_t \rho = c\Delta\rho$ on $[-a/2, a/2]^d$ with initial data μ_0 and Dirichlet boundary conditions. Let ϕ be a smooth test function on $\mathbb{R}^d$ with support in Λ. The quantity $M_\varepsilon(t) = \exp Z_\varepsilon(t)$, where

$$\begin{aligned} Z_\varepsilon(t) = {}& \varepsilon^{-d} \int \phi(x)[\rho(t, x) - \rho(0, x)]\, dx \\ & - c\varepsilon^{-2-d} \int_0^t \int \rho(s, x) \sum_e e^{\phi(\varepsilon(x+e)) - \phi(\varepsilon x)}\, dx\, ds \end{aligned}$$

is a martingale under P_ε and hence has expectation 1. By standard arguments (see [10]), one obtains from this the large deviation lower bound

$$\liminf_{k\to\infty} \varepsilon_k^d \log P_{\varepsilon_k}(B) \ge - \inf_{\rho(\cdot,\,\cdot)\in B} \int dt \|\partial_t \rho - c\Delta\rho\|^2_{-1,\, \rho^{-1}(t)},$$

where $\|f\|^2_{-1,\, \rho^{-1}} = \sup_\phi \{\int f\phi\, dx - \int |\nabla\phi|^2 \rho\, dx\}$ and B is any open set of time-dependent density functions. We conclude from this, and from the previous computation, that

$$\inf_{\rho(\cdot,\,\cdot)\in A} \int dt \|\partial_t \rho - c\Delta\rho\|^2_{-1,\, \rho^{-1}(t)} > 0.$$

This is really done in two steps. First we enlarge A to $\widetilde{A}$ slightly by considering density functions with initial data in a slightly larger set than $[-a/4, a/4]^d$ and which vanish at time a^2 on some subset $H \subset [-a/2, a/2]^d$ of Lebesgue measure slightly smaller than $\kappa_1 a^d$. Repeating the previous arguments, we obtain the lower bound above for the infimum over the interior of $\widetilde{A}$, which contains A.

In the general case $g(n) \le cn$, $M_\varepsilon(t)$ is a submartingale. Its expectation is therefore bounded by 1, and in particular we have

$$\limsup_{k\to\infty} \varepsilon_k^d \log E^{P_{\varepsilon_k}}[\exp Z_\varepsilon(t)] \le 0.$$

By standard arguments (see [10]), we can conclude from this the (nonoptimal) large deviation upper bound

$$\limsup_{k\to\infty} \varepsilon_k^d \log P_{\varepsilon_k}(A) \le -\inf_{\rho(\cdot,\cdot)\in A} \int dt \|\partial_t \rho - c\Delta\rho\|^2_{-1,\rho^{-1}(t)},$$

which is strictly negative by the previous argument. $\square$

PROOF OF THEOREM 4.1. A free particle here will mean one at a site where there are more than α particles and the number of holes at a site x is $\alpha - \eta_x$ if $\eta_x \le \alpha$.

Let t be a fixed time and let $\gamma > 0$ be arbitrary. Choose $\delta > 0$ and $\tau > 0$ and let $\mathscr{O}_{\delta,\tau} \subset \mathbb{R}^d$ be the set where the solution of the Stefan problem (1.7) is larger than $\alpha + 2^{d+1}\delta$ at time $t-\tau$. Note that from the definition of $\mathscr{A}_t$, we have $\mathscr{O}_{\delta,\tau} \to \mathscr{A}_t$ as $\tau, \delta \to 0$. Divide the set $\mathscr{O}_{\delta,\tau}$ into boxes of side length $a = \sqrt{\tau}$, throwing away those that are not completely contained in $\mathscr{O}_{\delta,\tau}$. It suffices to show that the probability that the density of holes at time t on a fixed such box is less than γ approaches 1 as ε goes to zero.

Let n be a positive integer. From the hydrodynamic limit we know that there are at least $\delta(a\varepsilon^{-1})^d$ free particles in a smaller box of side length $a/2$ with the same center as our fixed box at each time $t_0, \dots, t_{n-1}$, where $t_i = t - [(n-i)/n]\tau$, with probability which goes to 1 as $\varepsilon \to 0$. The positive integer n will be chosen large but bounded independent of ε later in the proof.

We study how many holes are removed in our fixed box in the time interval $[t_i, t_{i+1}]$. Our box has side length $L = a\varepsilon^{-1}$ sites and we know there are $N = \delta L^d$ free particles in a concentric box of side length $L/2$. We can pretend that only these N particles are able to remove holes, and furthermore that any free particle which leaves the box in which it started is annihilated immediately, for the true number of holes which is destroyed is only decreased in this way. Furthermore, we can pretend that our N particles only interact with each other and not with any other stray particles: let ζ_x denote the total number of free particles at x and let ξ_x denote the subset of these that come from our original group of N. If the ξ_x particles each jump at rate $g(\xi_x)/\xi_x$ and the other $\zeta_x - \xi_x$ each jump at rate $(g(\zeta_x) - g(\xi_x))/(\zeta_x - \xi_x)$, then the total jump rate is $g(\zeta_x)$ as required and the first N only interact among themselves. (Note that this relabeling is just the standard first-class/second-class picture and depends on the attractiveness of the system.) Let M denote the number of holes in our box (counted according to multiplicity). We also need only consider the case. $M > \gamma L^d$, for otherwise there is nothing to prove.

To summarize: each of our N particles moves according to a random walk with zero-range jump rates in the time interval $[t_i, t_{i+1}]$ until either (1) it hits a hole, in which case the particle is removed and the hole multiplicity at that site is reduced by one, or (2) it reaches the boundary of the larger box, in which case it is also removed from the system. The initial conditions consist of N, M, the positions of the N particles and the position of the M holes. These will be fixed for the estimates which we now derive, which are very rough and hold uniformly over all initial conditions, as well as all environments for the particles.

For $c_0 > 0$ with $c_0 N \le M$, let G denote the event that fewer than $c_0 N$ holes are filled during the specified time interval. If G happens, then there must exist a subset H of the original M holes, with cardinality less than $c_0 N$, so that all holes in the complement H^C are unfilled, and there must exist a subset W of the original N particles, with cardinality less than $c_0 N$, so that all particles in W^C avoid all the holes in H^C.

Now, let us choose any two subsets H and W of the original holes and particles, of cardinalities less than $c_0 N$. By the previous estimate, the probability that all particles in W^C avoid all the holes in H^C is bounded above by $e^{-c_1 N}$ for some $c_1 > 0$ for sufficiently small ε.

The number of possible choices for H and W is $\sum_{i,\,j=1}^{c_0 N-1}\binom{M}{i}\binom{N}{j}$. Using that $M \le \alpha L^d$ and the bound $\binom{n}{k} \le \exp\{2k\log(n/k)\}$ for large n and k with $k < n/2$, we can bound this by $N^2\exp\{c_2 N\}$, where $c_2 = 2c_0\log(\alpha/(c_0^2\delta))$. Choosing c_0 sufficiently small, we can bound this in turn by $\exp\{c_1 N/2\}$. Hence we have shown that there are constants $c_0 > 0$ and $c_3 = c_1/2$ independent of the environment, the initial conditions and ε such that

$$P(G) \le \exp\{-c_3\delta L^d\}.$$

Now we choose n large enough that $nc_0 N > M$, or in particular $n > \alpha/\delta c_0$ and we repeat the process in each of our time intervals, stopping if we ever have $M \le \gamma L^d$, and we say we are successful in the interval if G does not happen; in other words, if the number of holes filled is greater than or equal to $c_0 N$, or if we have stopped already. If we have success in every single interval, then for sure we have density less than γ in our box at time t. But by the previous estimate, the probability of success in every single interval is at least

$$\left(1-\exp\{-c_3\delta L^d\}\right)^n.$$

Since $L = a\varepsilon^{-1}$ and a, c_3 and n are independent of ε, this goes to 1 as $\varepsilon \to 0$. □

We now turn to applying Theorem 4.1 to obtain the desired shape theorems. Throughout the rest of this section, we will assume that the creation is at the origin only, at constant rate $c_1 = 1$. Recall that $A_t^\varepsilon \subset \varepsilon\mathbb{Z}^d$ is the occupied set in the diffusively scaled microscopic model and $\mathscr{A}_t = \{\mathbf{x}\colon \rho(\mathbf{x}, t) > 0\}$ is the occupied set in the Stefan problem (1.5). When $d = 2$, the creation and diffusion have the same scaling factor and so we obtain the following result.

COROLLARY 4.3. *Assume that* $d = 2$. *As* $t \to \infty$, $\operatorname{dist}\left(A_t^1/\sqrt{t}, \mathscr{A}_1\right) \to 0$ *in probability as* $t \to \infty$.

We now show that in $d = 2$, $\mathscr{A}_1$ is a circular disc with the conjectured radius by constructing an explicit radially symmetric solution of (1.5).

Putting the equation into polar coordinates, the problem is to show that there exists a constant $K \in (0, \infty)$, and a function $\rho(r, t)$ which is smooth in the interior of $\{(r, t)\colon r \le K\sqrt{t}\}$ and satisfies

$$\frac{\partial \rho}{\partial t} = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\gamma(\rho) + \frac{1}{2\pi r}\delta_0,$$

with boundary conditions

$$\rho(K\sqrt{t}, t) = 0,$$
$$\frac{\partial \rho}{\partial r}(K\sqrt{t}, t) = -\frac{K\alpha}{2\sqrt{t}\gamma'(0)}.$$

We look for self-similar solutions of the form $\rho(r, t) = v(rt^{-1/2})$. Writing $s = rt^{-1/2}$, $F = \gamma^{-1}$ and $u = \gamma(v)$, one gets

$$\begin{aligned} &u''(s) + \left(\frac{1}{s} + \frac{1}{2}sF'(u)(s)\right)u'(s) + \frac{1}{2\pi s}\delta_0 = 0, \\ &u(K) = 0, \qquad u'(K) = -\frac{1}{2}\alpha K. \end{aligned} \tag{4.1}$$

It follows that, off $s = 0$, $(su')' = -\frac{1}{2}s^2F'(u)u'$ and therefore,

$$su'(s) = -\frac{1}{2}\alpha K^2 + \int_s^K \frac{1}{2}\sigma^2 F'(u)(\sigma)u'(\sigma)\,d\sigma. \tag{4.2}$$

In addition, $u(K) = 0$ and $su'(s) \to -1/(2\pi)$ as $s \to 0$. Recall first that F' is bounded above and below by positive constants and note that (4.1) implies that $u' < 0$. Moreover, (4.2) and Gronwall's inequality imply that, for some constant C (which depends only on γ), $-\frac{1}{2}\alpha K^2 \exp(CK^2) \le su'(s) \le -\frac{1}{2}\alpha K^2$. As $su'(s)$ is also increasing on $(0, K)$, it has a limit as $s \to 0$, which varies continuously between 0 and $-\infty$ as K varies between 0 and ∞. Therefore, there exists a K such that $\lim_{s\to 0} su'(s) = -1/(2\pi)$, and therefore K can be chosen so that a solution of (4.1) exists [such a solution must be unique, since it produces a solution of (1.5)].

When $\gamma(u) = u$, and therefore $F'(u) = 1$, (4.1) can be explicitly solved, giving

$$u(s) = \int_s^K \frac{1}{2\pi s}\exp(-s^2/4)\,ds,$$

which forces $K = K(\alpha)$ to be the solution of

$$\exp(-K^2/4) = \pi\alpha K^2. \tag{4.3}$$

For example, to two significant digits, $K = 0.54$ if $\alpha = 1$, $K = 0.39$ if $\alpha = 2$ and $K \sim 1/\sqrt{\pi\alpha}$ as $\alpha \to \infty$. It is amusing to note that, for $\alpha = 1$, the upper bound on K obtained by pretending that all particles are killed is $1/\sqrt{\pi} \approx 0.56$. Thus the system conspires to keep a particular fraction ($\approx$7%) of the particles alive.

In particular, Lemma 4.2 now implies the following shape theorem.

THEOREM 4.4. *In dimension $d = 2$, there exists a constant K, such that*

$$\mathrm{dist}(A_1(t)/\sqrt{t}, B(0, K)) \to 0$$

as $t \to \infty$. In the linear case $\gamma(u) = u$, K is given by (4.3).

Now we address the long-term behavior of both the particle system and the Stefan problem in dimensions other than 2. For constant creations at a single site, there are now no self-similar solutions. One way to circumvent this problem is to make the creation rate variable with time. Assume that the creation rate at the origin c_1 is a function of t and satisfies $c_1(t)t^{-(d-2)/2} \to 1$ as $t \to \infty$. Then, after speeding up time by factor ε^{-2}, and sending $\varepsilon \to 0$, the macroscopic equation becomes

$$\frac{\partial \rho}{\partial t} = \Delta\gamma(\rho) + t^{(d-2)/2}\delta_0,$$

or in polar coordinates,

$$\frac{\partial \rho}{\partial t} = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\gamma(\rho) + \frac{\Gamma(d/2)}{2\pi^{d/2}} r^{1-d} t^{(d-2)/2}\delta_0,$$

where Γ is the gamma function. The convergence is proved in the same way as for constant creation rate. A self-similar solution $\rho = v(rt^{-1/2})$ exists for any γ (with the same proof as before). In the case $\gamma(u) = u$,

$$v(s) = \frac{\Gamma(d/2)}{2\pi^{d/2}} \int_s^K s^{1-d} \exp(-s^2/4)\, ds, \quad \text{where} \quad \frac{\Gamma(d/2)}{\pi^{d/2}} \exp(-K^2/4) = \alpha K^d.$$

Now we study the problem of long time behavior in dimensions different from 2 in the case of constant creation rate at the origin. It is proved in [3] that when $d \geq 3$, $\gamma(u) = u$ and $\alpha = 1$, the asymptotic behavior of A_t^1 is the same as when the source is discrete; that is, with probability 1, for every $\delta > 0$,

$$((1-\delta)t^{1/d} \cdot B(0, K)) \cap \mathbb{Z}^d \subset A_t^1 \subset (1+\delta)t^{1/d} \cdot B(0, K), \tag{4.4}$$

eventually in t. Here, $K^d = (d\Gamma(d/2))/2\pi^{d/2}$, so that $B(0, K)$ has unit volume. Our first theorem shows that the occupied set in the Stefan problem expands at the same rate.

PROPOSITION 4.5. *Assume that $d \geq 3$, $\gamma(\rho) = \rho$, $\alpha = 1$ and $\delta > 0$ is fixed. For sufficiently large t,*

$$(1-\delta)t^{1/d} \cdot B(0, K) \subset \mathscr{A}_t \subset (1+\delta)t^{1/d} \cdot B(0, K),$$

with the same K as in (4.4). *In the case of general γ, there exist two constants $0 < K_1 < K_2 < \infty$, so that*

$$t^{1/d} \cdot B(0, K_1) \subset \mathscr{A}_t \subset t^{1/d} \cdot B(0, K_2).$$

PROOF. The idea is to construct sub- and supersolutions of the Stefan problem (1.5) out of the explicit solution

$$v(r,t) = \int_0^t (4\pi s)^{-d/2} \exp(-r^2/4s)\, ds = \frac{1}{4\pi^{d/2}} r^{2-d} \int_{r^2/4t}^{\infty} s^{d/2-2} e^{-s}\, ds$$

of the heat equation $\partial v/\partial t = \Delta v + \delta_0$. Consider first the case $\gamma(\rho) = \rho$. We have

$$\begin{aligned}\frac{\partial v}{\partial r}(r,t) &= -\frac{r^{1-d}}{2\pi^{d/2}} \int_{r^2/4t}^{\infty} s^{(d-2)/2} e^{-s}\, ds \\ &= -\frac{r^{1-d}}{2\pi^{d/2}} \big(\Gamma(d/2) + O(r^d t^{-d/2})\big).\end{aligned} \tag{4.5}$$

Let us take a ball of radius $Ct^{1/d}$. The rate of expansion of the boundary is $Ct^{1/d-1}/d$. Note that this coincides with $[-(\partial v)/(\partial r)](Ct^{1/d}, t)$ up to order $t^{1/d-d/2}$ exactly when $C = K$, where $K^d = (d\Gamma(d/2))/2\pi^{d/2}$ as in (4.4), that is, exactly when the ball has volume t. Define

$$\rho(r,t) = v(r,t) - v(Ct^{1/d}, t)$$

for $r \le Ct^{1/d}$ and $v = 0$ outside the ball. If $C > K$, then, after a certain time, ρ is a supersolution of the Stefan problem, and if $C < K$, then it is eventually a subsolution. The proposition then follows from Lemma 2.9.

For the case of general γ, note that the spatial coupling term $\Delta\gamma(\rho)$ can be written as $\nabla \cdot D\nabla$, where from the assumptions, $\gamma' = D$ is uniformly bounded above and below away from zero. Aronson's estimates [6] tell us that there are positive finite constants c_1, c_2, b_1, b_2 such that

$$\begin{aligned}&c_1(4\pi s)^{-d/2} \exp(-b_1|\mathbf{x}-\mathbf{y}|^2/4(t-s)) \\ &\qquad \le p(s,\mathbf{x},t,\mathbf{y}) \le c_2(4\pi s)^{-d/2} \exp(-b_2|\mathbf{x}-\mathbf{y}|^2/4(t-s)),\end{aligned}$$

where $p(s,\mathbf{x},t,\mathbf{y})$ is the solution at $(\mathbf{y}, t)$ of the corresponding heat equation, starting from a Dirac mass at $\mathbf{x}$ at time s. Therefore, the flux across the boundary of the ball of radius $Ct^{1/d}$ is bounded above and below in terms of the corresponding flux in the linear case (4.5). But now we can repeat the same argument as in the linear case to obtain the upper and lower bounds. □

In one dimension, the proportion of created particles which are killed converges to 0 as time progresses. Therefore, as pointed out in [3], the rate of boundary expansion is of much slower order than t. Again, [3] have a conjecture in the linear case which we now state and prove as a theorem.

THEOREM 4.6. *Assume that* $d = 1$, $\gamma(u) = u$ *and* $\alpha = 1$. *Then for every* $\varepsilon > 0$, *with probability* 1, *for sufficiently large* t,

$$(1-\varepsilon)\sqrt{2t\log t} \cdot [-1,1] \cap \mathbb{Z} \subset A_t^1 \subset (1+\varepsilon)\sqrt{2t\log t} \cdot [-1,1]. \tag{4.6}$$

PROOF. Choose $1 < c_1 < c < 1+\varepsilon$. Let N_t be the number of particles created during the time interval $[0,t]$ which exit $[-c\sqrt{2t\log t}, c\sqrt{2t\log t}]$ before time t. Moreover, let $p_{s,t}$ be the probability of the event that a random walk started at the origin at time $s \in [0,t]$ visits a site outside $[-c\sqrt{2t\log t}, c\sqrt{2t\log t}]$ during the time interval $[s,t]$. Observe first that $p_{s,t} \le p_{0,t}$ by a simple coupling. Next, let S_n be the discrete-time simple symmetric random walk started at 0, and W_t the standard Brownian motion. Then, by elementary large deviations,

$$p_{0,t} \le 2P(\max\{S_k: 0 \le k \le 2c_1 t\} > c\sqrt{2t\log t}) + e^{-\beta t},$$

for some $\beta > 0$. Using first the Skorohod embedding, and then the reflection principle, we obtain

$$\begin{aligned} p_{0,t} &\le 2P\big(\max\{W_s: 0 \le t \le 2c_1 t\} > c\sqrt{2t\log t}\big) + e^{-\beta t} \\ &\le 4P\big(W_{2c_1 t} > c\sqrt{2t\log t}\big) + e^{-\beta t} \\ &\le C\cdot t^{-c^2/(2c_1)} < C\cdot t^{-c/2}, \end{aligned}$$

for large t. Let N'_t be the number of particles created during the time interval $[0,t]$. Then

$$P(N_t \ge \sqrt{t}) \le P(N_t \ge \sqrt{t}, N'_t \le 2t) + P\big(N'_t > 2t\big).$$

Here, $P\big(N'_t > 2t\big)$ is exponentially small and

$$\begin{aligned} P\big(N_t \ge \sqrt{t}, N'_t \le 2t\big) &\le \exp\big(-\sqrt{t}\big)E\big(\exp(N_t)\mathbf{1}_{N'_t \le 2t}\big) \\ &\le \exp\big(-\sqrt{t}\big)\big(1+2p_{0,t}\big)^{2t} \\ &\le \exp\big(-\sqrt{t} + Ct^{1-c/2}\big). \end{aligned}$$

Thus, by the Borel–Cantelli lemma, $P(N_t \le \sqrt{t}\text{ eventually}) = 1$. Hence, even if all of N_t particles get killed outside $[-c\sqrt{2t\log t}, c\sqrt{2t\log t}]$ by time t, A_t^1 is with probability 1 eventually included in $[-c\sqrt{2t\log t} - 2\sqrt{t}, c\sqrt{2t\log t} + 2\sqrt{t}]$. The upper bound in (4.7) now follows since $(1+\varepsilon-c)\sqrt{2t\log t}$ eventually exceeds $2\sqrt{t}$.

To prove the lower bound, let $A_t^1 = [-L_t, R_t]$. We will first show that, with large probability, at least one of L_t and R_t must be at least $c\sqrt{2t\log t}$, for some small c. Then we will show that both L_t and R_t must satisfy this lower bound; in the final step we will bootstrap c close to 1.

To this end, fix $c < 1-2\varepsilon$ and let $c_1 = c/(1-2\varepsilon)$. Let us momentarily consider the free particle system with creation only, without killing. In this case, let N_t (resp. N'_t) be the number of particles created in $[0,\varepsilon t]$, whose position is at least $c\sqrt{2t\log t}$ (resp. at most $-c\sqrt{2t\log t}$) at time t. Again, $P(N_t \le c\sqrt{2t\log t}) \le E(\exp(c\sqrt{2t\log t} - N_t))$, and normal approximation can be used to show that

$$E(\exp(-N_t)) \le \exp\big(-C\varepsilon t^{1-c_1^2/2}\big), \tag{4.8}$$

so that $P\big(N_t \geq c\sqrt{2t\log t}$ eventually$) = 1$ and by symmetry the same is true for N_t replaced by N'_t. The natural coupling between the internal DLA and the free particle system is the one in which the particles in the internal DLA are tagged but allowed to continue to diffuse. Under this coupling, if N_t and N'_t are both larger than $c\sqrt{2t\log t}$, then $R_t \vee L_t \geq c\sqrt{2t\log t}$.

From the previous paragraph, it follows that for large enough t, $R_{\varepsilon t} \vee L_{\varepsilon t} \geq c\varepsilon\sqrt{2t\log t}$. Assume, for example, that $L_{\varepsilon t} \geq c\varepsilon\sqrt{2t\log t}$. Now make particles created *after* time εt execute free random walks, and define N_t^1 to be the number of particles created in $[\varepsilon t, 2\varepsilon t]$, which satisfy the following two conditions:

1. the particle's position at time t is at least $c\sqrt{2t\log t}$, and
2. the particle does not go below $-L_{\varepsilon t}$ during the time interval $[\varepsilon t, t]$.

Once again, one can use the Skorohod embedding and the reflection principle for Brownian motion to show that a single particle will satisfy (1) and (2) above with probability at least $\varepsilon t^{-c_1^2/2}$, and so (4.8) holds with N_t replaced by N_t^1 and ε replaced by ε^2. Thus it follows that, with probability 1, for large t, $R_t \wedge L_t \geq c\varepsilon\sqrt{2t\log t}$.

For the final bootstrapping step, notice that what we proved so far implies that $R_{\varepsilon t} \wedge L_{\varepsilon t} \geq c\varepsilon^2\sqrt{t\log t}$ for t large enough. Then repeat the argument in the previous paragraph with ε replaced by ε^2 to show that $R_t \wedge L_t \geq c\sqrt{2t\log t}$ with probability at least $\exp(-C\varepsilon^3 t^{1-c_1^2/2})$. □

Again, it turns out that $\mathscr{A}_t$ has exactly the same rate of expansion.

PROPOSITION 4.7. *Assume that* $d = 1$, $\gamma(v) = v$ *and* $\alpha = 1$, *and fix an* $\varepsilon > 0$. *Then*, *for large* t,

$$(1-\varepsilon)\sqrt{2t\log t}\cdot[-1,1] \subset \mathscr{A}_t \subset (1+\varepsilon)\sqrt{2t\log t}\cdot[-1,1].$$

In the case of general γ, *there exist two constants* $0 < K_1 < K_2 < \infty$, *so that* $K_1\sqrt{t\log t}\cdot[-1,1] \subset \mathscr{A}_t \subset K_2\sqrt{t\log t}\cdot[-1,1]$.

PROOF. With notation as in the proof of Proposition (4.5), we need to prove that for large t,

$$-\frac{\partial v}{\partial r}\big(C\sqrt{2t\log t}, t\big) \geq C\sqrt{\frac{\log t}{2t}}, \tag{4.7}$$

when $C < 1$, while the inequality is reversed when $C > 1$. This follows from the fact that the integral on the left of (4.7) is asymptotic to a constant multiple of $t^{-C/2}(\log t)^{-1/2}$. The case of general γ again follows as in the proof of Proposition (4.5) by using Aronson's estimates. □

REMARK. Theorem 4.6, as well as the upper and lower bounds obtained in Propositions 4.6 and 4.7, can easily be extended to the case of multiple creation sites or $\alpha > 1$, thereby proving the shape theorems in these cases as well.

5. Invariant measures. In this section we prove Lemma 3.3, which identifies the infinite volume invariant measures of our degenerate zero-range processes. We use the methods of [1], [13], with appropriate modifications to deal with the degeneracy.

The family of measures μ^β on configurations $\eta \in \{0, 1, \ldots\}^{\mathbb{Z}^d}$ comprises the following two sets of measures:

1. for every $u \geq \alpha$, the translation-invariant product measures in which, for every x, $\eta_x \geq \alpha$ and $\zeta_x = \eta_x - \alpha$ is distributed as the marginal for the grand canonical invariant measure for the zero-range dynamics L_1 from (1.2) with density $\rho = u - \alpha$, and
2. all Dirac masses on configurations in $\{0, 1, \ldots, \alpha\}^{\mathbb{Z}^d}$.

It is clear that any such μ^β is invariant. What we need is the converse.

THEOREM 5.1. *Any invariant, translation-invariant measure for L_0 defined in* (1.1) *is a mixture of μ^β. More precisely, any extremal invariant, translation-invariant measure is one of μ^β.*

The only methods available at the present time to prove such a result rely on either attractiveness.or duality. Since there is no useful duality available in our context, we have to assume attractiveness and use a method based on coupling. The coupled process makes the usual attempt to "move the two processes in unison as much as possible." Formally, the generator of the coupled process is given by

$$\begin{aligned}\overline{L} f(\eta, \xi) = \sum_{x \sim y} & \mathbf{1}_{\eta_x > \xi_x} (a(\eta_x) - a(\xi_x))(f(\eta^{xy}, \xi) - f(\eta, \xi)) \\ & + \mathbf{1}_{\eta_x < \xi_x} (a(\xi_x) - a(\eta_x))(f(\eta, \xi^{xy}) - f(\eta, \xi)) \\ & + a(\eta_x \wedge \xi_x)(f(\eta^{xy}, \xi^{xy}) - f(\eta, \xi)).\end{aligned}$$

Note that we do not really need to require the rates of the zero-range process performed by the particles on top of the occupied set to be strictly increasing for the arguments in this section to hold. This assumption was made to minimize technical difficulties elsewhere in the proof.

In the statements of Lemmas 5.2–5.5, we assume that $\bar{\nu}$ is an invariant, translation-invariant measure for the coupled process.

LEMMA 5.2. *For any two neighboring sites x_0 and y_0,*

$$\bar{\nu}\{(\eta, \xi) \colon \eta_{x_0} < \xi_{x_0}, \xi_{y_0} < \eta_{y_0}, \alpha < \eta_{y_0}\} = 0. \tag{5.1}$$

PROOF. Let $f(\eta, \xi) = (\xi_{x_0} - \eta_{x_0})_+$. Then

$$\begin{aligned}\overline{L} f = \sum_{x \sim x_0} & \mathbf{1}_{\eta_{x_0} > \xi_{x_0}} (a(\eta_{x_0}) - a(\xi_{x_0}))(f(\eta^{x_0 x}, \xi) - f(\eta, \xi)) \\ & + \mathbf{1}_{\eta_{x_0} < \xi_{x_0}} (a(\xi_{x_0}) - a(\eta_{x_0}))(f(\eta, \xi^{x_0 x}) - f(\eta, \xi))\end{aligned}$$

$$
\begin{aligned}
&+ a(\eta_{x_0} \wedge \xi_{x_0})(f(\eta^{x_0 x}, \xi^{x_0 x}) - f(\eta, \xi)) \\
&+ \mathbf{1}_{\eta_x > \xi_x}(a(\eta_x) - a(\xi_x))(f(\eta^{x x_0}, \xi) - f(\eta, \xi)) \\
&+ \mathbf{1}_{\eta_x < \xi_x}(a(\xi_x) - a(\eta_x))(f(\eta, \xi^{x x_0}) - f(\eta, \xi)) \\
&+ a(\eta_x \wedge \xi_x)(f(\eta^{x x_0}, \xi^{x x_0}) - f(\eta, \xi)) \\
= \sum_{x \sim x_0} & 0 \\
\text{(I)} \quad &+ \mathbf{1}_{\xi_{x_0} > \eta_{x_0}} - (a(\xi_{x_0}) - a(\eta_{x_0}))(-1) \\
&+ 0 \\
\text{(II)} \quad &+ \mathbf{1}_{\eta_x > \xi_x, \xi_{x_0} > \eta_{x_0}}(a(\eta_x) - a(\xi_x))(-1) \\
\text{(III)} \quad &+ \mathbf{1}_{\eta_x < \xi_x, \xi_{x_0} \geq \eta_{x_0}}(a(\xi_x) - a(\eta_x))(+1) \\
&+ 0,
\end{aligned}
$$

where $\sim$ denotes nearest neighbors on the lattice. Then

$$E_{\bar{\nu}}\Bigg(\sum_{x \sim x_0} \text{expression (I)} \Bigg) = -\sum_{x \sim x_0} \sum_{k > \ell} (a(k) - a(\ell)) \bar{\nu}(\xi_{x_0} = k, \eta_{x_0} = \ell),$$

while

$$E_{\bar{\nu}}\Bigg(\sum_{x \sim x_0} \text{expression (III)} \Bigg) = \sum_{x \sim x_0} \sum_{k > \ell} (a(k) - a(\ell)) \bar{\nu}(\xi_x = k, \eta_x = \ell, \xi_{x_0} \geq \eta_{x_0}).$$

Therefore, by translation invariance,

$$
\begin{aligned}
E_{\bar{\nu}}(\overline{L} f) &\leq E_{\bar{\nu}}\Bigg(\sum_{x \sim x_0} \text{expression (II)} \Bigg) \\
&\leq -\sum_{k < \ell} (a(\ell) - a(k)) \bar{\nu}(\xi_{y_0} = k, \eta_{y_0} = \ell, \xi_{x_0} > \eta_{x_0}).
\end{aligned}
$$

However, as $\bar{\nu}$ is invariant, $E_{\bar{\nu}}(\overline{L} f) = 0$. Since the rates $a(\ell)$ are strictly increasing for $\ell > \alpha$, $\nu(\xi_{y_0} = k, \eta_{y_0} = \ell, \xi_{x_0} > \eta_{x_0}) = 0$ whenever $k < \ell$ and $\alpha < \ell$, which is equivalent to the statement of the lemma. $\square$

LEMMA 5.3. *For any two neighboring sites x_0 and y_0,*

$$\bar{\nu}\{(\eta, \xi)\colon \eta_{y_0} > \alpha, \eta_{x_0} < \alpha\} = 0. \tag{5.2}$$

PROOF. This time, let $f = (\eta_{x_0} - \alpha)_+$. Then

$$\overline{L} f = \sum_{x \sim x_0} \big(a(\eta_{x_0}) \mathbf{1}_{\eta_{x_0} > \alpha}(-1) + a(\eta_x) \mathbf{1}_{\eta_{x_0} \geq \alpha}(+1) \big).$$

Therefore,

$$\begin{aligned} E_{\bar\nu}(\overline{L}f) &= \sum_{x\sim x_0}\sum_{k>\alpha} a(k)\big(-\bar\nu(\eta_{x_0}=k)+\bar\nu(\eta_x=k,\eta_{x_0}\ge\alpha)\big)\\ &= -\sum_{x\sim x_0}\sum_{k>\alpha} a(k)\bar\nu(\eta_x=k,\eta_{x_0}<\alpha). \end{aligned}$$

This implies that, for each $k>\alpha$, $\bar\nu(\eta_{y_0}=k,\eta_{x_0}<\alpha)=0$. $\square$

We will shorten $\max\eta=\max\{\eta_x\colon x\in\mathbb{Z}^d\}$ and $\min\eta=\min\{\eta_x\colon x\in\mathbb{Z}^d\}$.

LEMMA 5.4. $\bar\nu(\max\eta>\alpha,\min\eta<\alpha)=0$.

PROOF. It suffices to show that (5.2) holds for arbitrary sites x_0 and y_0. To this end we will prove, by induction on $k\ge 1$, that (5.2) holds for all x_0, y_0 for which $\|x_0-y_0\|_1=k$, where $\|x\|_1=\sum_{i=1}^d|x_i|$ is the L^1 distance on the lattice. To prove the $k-1\to k$ step, take a y_1 with $\|x_0-y_1\|_1=k-1$ and $\|y_1-y_0\|_1=1$. Choose the function $f=\mathbf{1}_{\eta_{y_1}>\alpha,\eta_{x_0}<\alpha}$. Then, by the induction hypothesis, $f=0$ $\bar\nu$-a.s., and therefore, $\overline{L}f\ge 0$ $\bar\nu$-a.s. However, $E_{\bar\nu}(\overline{L}f)=0$, thus so is $E_{\bar\nu}$ of the part of $\overline{L}f$ which corresponds to a particle jumping from y_0 to y_1 in η. That is,

$$\begin{aligned} 0 &= E_{\bar\nu}\big(a(\eta_{y_0})\mathbf{1}_{\eta_{y_1}\ge\alpha,\eta_{x_0}<\alpha}\big)\\ &= \sum_{k>\alpha} a(k)\bar\nu\big(\eta_{y_0}=k,\eta_{y_1}\ge\alpha,\eta_{x_0}<\alpha\big)\\ &= \sum_{k>\alpha} a(k)\bar\nu\big(\eta_{y_0}=k,\eta_{x_0}<\alpha\big), \end{aligned}$$

the last inequality being the consequence of Lemma 5.3 applied to y_0 and y_1. Thus $\bar\nu(\eta_{y_0}=k,\eta_{x_0}<\alpha)=0$ for any $k>\alpha$. $\square$

LEMMA 5.5. *The two marginals of $\bar\nu$ are monotonely coupled, that is*,

$$\begin{aligned} \bar\nu\Big(\Big(&\cap_{x\in\mathbb{Z}^d}\big\{(\eta_x-\alpha)_+\ge(\xi_x-\alpha)_+\big\}\Big)\\ &\cup\Big(\cap_{x\in\mathbb{Z}^d}\big\{(\eta_x-\alpha)_+\le(\xi_x-\alpha)_+\big\}\Big)\Big)=1. \end{aligned}$$

PROOF. This proof is very similar to the previous one, so we just point out the main steps. It suffices to prove that (5.1) holds for arbitrary sites x_0 and y_0, that is, for $\|x_0-y_0\|_1=k$ for $k\ge 1$. Again, take a y_1 with $\|x_0-y_1\|_1=k-1$ and $\|_1-y_0\|=1$, and choose the function $f=\mathbf{1}_{\eta_{y_1}>\xi_{y_1},\xi_{x_0}>\eta_{x_0},\alpha<\eta_{y_1}}$. As before, the induction hypothesis implies that $E_{\bar\nu}$ of the part of $\overline{L}f$ which corresponds to a particle jumping from y_0 to y_1 in η, but not in ξ, is 0:

$$\begin{aligned}0 &= E_{\bar\nu}\big((a(\eta_{y_0}) - a(\xi_{y_0}))\mathbf{1}_{\eta_{y_0}>\xi_{y_0}} f(\eta^{y_0y_1}, \xi)\big)\\ &= \sum_{k>\ell}(a(k)-a(\ell))\bar\nu(\eta_{y_0}=k, \xi_{y_0}=\ell, \eta_{y_1}\ge \xi_{y_1}, \xi_{x_0} > \eta_{x_0}, \alpha \le \eta_{y_1})\\ &= \sum_{k>\ell}(a(k)-a(\ell))\bar\nu(\eta_{y_0}=k, \xi_{y_0}=\ell, \xi_{x_0} > \eta_{x_0}, \alpha \le \eta_{y_1}),\end{aligned}$$

by Lemma 5.3. It follows that

$$\begin{aligned}0 &= \bar\nu(\eta_{y_0} > \xi_{y_0}, \eta_{y_0} > \alpha, \xi_{x_0} > \eta_{x_0}, \alpha \le \eta_{y_1})\\ &= \bar\nu(\eta_{y_0} > \xi_{y_0}, \eta_{y_0} > \alpha, \xi_{x_0} > \eta_{x_0}),\end{aligned}$$

by Lemma 5.3. This completes the proof. □

PROOF OF THEOREM 5.1. Assume that μ_1 is an extremal invariant translation-invariant measure for the process. Pick another extremal invariant translation-invariant measure μ_2 (to be specified later). Then there exists an extremal invariant translation-invariant measure $\bar\nu$ for the coupled process, with marginals μ_1 and μ_2 (Lemma 4.3 in [1]).

By Lemma 5.4, one can decompose the measure μ_1 as follows:

$$\begin{aligned}\mu_1(\eta\in\cdot) &= \mu_1(\eta\in\cdot\,|\min\eta<\alpha)\mu_1(\min\eta<\alpha)\\ &\quad+\mu_1(\eta\in\cdot\,|\max\eta>\alpha)\mu_1(\max\eta>\alpha)\\ &\quad+\mu_1(\eta\in\cdot\,|\eta\equiv\alpha)\mu_1(\eta\equiv\alpha)\\ &= \mu_1(\eta\in\cdot\,|\max\eta\le\alpha\cdot\eta\not\equiv\alpha)\mu_1(\min\eta<\alpha)\\ &\quad+\mu_1(\eta\in\cdot\,|\min\eta\ge\alpha, \eta\not\equiv\alpha)\mu_1(\max\eta>\alpha)\\ &\quad+\mu_1(\eta\in\cdot\,|\eta\equiv\alpha)\mu_1(\eta\equiv\alpha).\end{aligned}$$

The three conditional measures above are invariant and translation invariant. Since μ_1 is extremal, one of the three nonconditional probabilities above must be 1. If either $\mu_1(\min\eta<\alpha)=1$ or $\mu_1(\eta\equiv\alpha)=1$, then $\mu_1(\max\eta\le\alpha)=1$, so there is nothing to prove. If $\mu_1(\max\eta>\alpha)=1$, then $\eta_x\ge\alpha$ for every x μ_1-a.s. Assume this for the rest of the proof.

Now let μ_1 be an extremal invariant measure with $u=E_{\mu_1}(\eta_0)\ge\alpha$. Moreover, let μ_2 be the extremal measure μ_u from subset (1) of the family μ^β; that is, μ_u is the product measure in which, for every x, $\eta_x\ge\alpha$ and $\zeta_x=\eta_x-\alpha$ has the marginal distribution of the grand canonical invariant measure for the nondegenerate zero-range process L_1 defined in (1.2). Then, by Lemma 5.5 and the fact that $\bar\nu$ is extremal (see the proof of Lemma 4.5 in [1]).

$$\bar\nu\big(\cap_{x\in\mathbb{Z}^d}\{\eta_x\ge\xi_x\}\big)=1 \qquad\text{or}\qquad \bar\nu\big(\cap_{x\in\mathbb{Z}^d}\{\eta_x\le\xi_x\}\big)=1. \tag{5.3}$$

As μ_1 and μ_2 have the same density, $E_{\bar\nu}(\eta_x-\xi_x)=0$ for every x. This forces $\eta_x=\xi_x$ $\bar\nu$-a.s. in either case of (5.3), that is, $\bar\nu(\{(\eta,\xi)\colon \eta\equiv\xi\})=1$. In particular, $\mu_1=\mu_u$. □

Notes added in proof. (i) Internal DLA models, and Stefan problems, can be classified, as subcritical, critical or supercritical according to whether

$$t^{-d/2}\int_0^t c(s)\,ds$$

goes to 0, remains bounded or goes to ∞, as t goes to ∞, where $c(t)$ is the (time-dependent) creation rate at the origin. In this scheme, [3] study the subcritical regime and we study the critical regime for internal DLA. The [3] method for continuous-time models is by comparison with discrete time. One can consider also the "soft" version of their models, say in $d \geq 3$ with random walk at rate ε^{-2} and creation at the origin at rate ε^{-d}. In this article, we proved the hydrodynamic limit for such models are Stefan problems and one can check easily that they are of subcritical type. Then one can reverse the comparison argument of [3] to infer the shape theorem for the original discrete- or continuous-time internal DLA from that of the "soft" version. In this way, the hydrodynamic limit method can imply shape theorems for the subcritical cases even though they have no hydrodynamic scaling themselves. We know of no results in the supercritical case except in one dimension.

(ii) Since this paper was submitted, related work has been done by BenArous and Ramirez and by Funaki. BenArous and Ramirez have an independent proof of the one-dimensional case using order statistics. They also study the large deviations of the lifetime of a tagged particle in the subcritical case with random obstacles: the first ξ_x particles at x are frozen, where ξ is a stationary random field. Funaki studies a model similar to ours without the creation.

(iii) After we completed this work, we were shown the article [4], which studies a model related to ours in one dimension, also making the connection with a Stefan problem. Unfortunately, some of the details of their argument remain unclear to us, in particular the crucial final sentence of the proof of their Proposition 2.1.

Acknowledgment. The authors would like to thank one of the referees for a very critical reading of the manuscript and many comments which led to the present improved version of the article.

REFERENCES

[1] ANDJEL, E. (1982). Invariant measures for the zero-range processes. *Ann. Probab*. **10** 525–547.

[2] BENAROUS, G. and RAMIREZ, A. Diffusion and saturation processes in random media. Preprint.

[3] BRAMSON, M., GRIFFEATH, D. and LAWLER, G. (1990). Internal diffusion limited aggregation. *Ann. Probab.* **20** 2117–2140.

[4] CHAYES, L. and SWINDLE, G. (1996). Hydrodynamic limits for one-dimensional particle systems with moving boundaries. *Ann. Probab.* **24** 559–598.

[5] CHANG, C. C. and YAU, H.-T. (1992). Fluctuations of one-dimensional Ginzburg–Landau models in nonequilibrium. *Comm. Math. Phys.* **145** 209–234.

[6] DAVIES, E. B. (1989). *Heat Kernels and Spectral Theory*. Cambridge Univ. Press.

[7] DIACONIS, P. and FULTON, W. (1991). A growth model, game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Semin. Mat. Univers. Politecn. Torino* **49** 95–119.
[8] FRIEDMAN, A. (1982). *Variational Principles and Free Boundary Problems*. Wiley-Interscience, New York.
[9] FUNAKI, T. Preprint.
[10] KIPNIS, C. and LANDIM, C. (1999). Scaling limits of interacting particle systems. *Grundlehren der Mathematischen Wissenschaften* **320**. Springer, Berlin.
[11] LANDIM, C., and SETHURAMAN, S. and VARADHAN, S. R. S. (1996). Spectral gap for zero-range processes. *Ann. Probab.* **24** 1871–1902.
[12] LAWLER, G. F. (1995). Subdiffusive fluctuations for internal diffusion limited aggregation. *Ann. Probab.* **23** 71–86.
[13] LIGGETT, T. (1976). Coupling the simple exclusion process. *Ann. Probab.* **4** 339–356.
[14] LIONS, J.-L. and MAGENES, E. (1972). Non-homogeneous boundary value problems and applications. *Grundlehren der Mathematischen Wissenschaften* **181–182**. Springer, Berlin.
[15] LUSTERNIK, L. A. and SOBOLEV, V. J. (1961). *Elements of Functional Analysis*. Gordon and Breach, New York.
[16] MEIRMANOV, A. M. (1992). *The Stefan Problem*. de Gruyter, Berlin.
[17] SPOHN, H. (1993). Interface motion in models with stochastic dynamics. *J. Statist. Phys.* **71**.
[18] YAU, H.-T. (1994). Metastability of Ginzburg–Landau model with a conservation law. *J. Statist. Phys.* **74** 705–742.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
DAVIS, CALIFORNIA 95616
E-MAIL: gravner@math.ucdavis.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF TORONTO
100 ST. GEORGE STREET
TORONTO, ONTARIO M5S 3G3
CANADA
E-MAIL: quastel@math.toronto.edu