

# Lectures on Dynamics in Models of Coarsening and Coagulation

Robert L. Pego

*Department of Mathematical Sciences  
Carnegie Mellon University  
Pittsburgh, PA 15213, USA  
E-mail: rpego@cmu.edu*

## Contents

1	Introduction	2
2	A hierarchy of domain coarsening models in one space dimension	5
2.1	Domain walls in the Allen-Cahn equation	5
2.2	Domain wall dynamics by restricted gradient flow	7
2.3	Punctuated equilibrium and 1D bubble bath	11
2.4	Mean-field model of domain growth — The Gally-Mielke transform	13
2.5	Proof of universal self-similar behavior	19
3	Models of domain coarsening in two and three dimensions	23
3.1	Diffuse and sharp-interface models of nanoscale island coarsening	24
3.2	Gradient structure for Mullins-Sekerka flow	31
3.3	Monopole models by restricted gradient flow of surface energy	33
3.4	Lifshitz-Slyozov-Wagner mean-field model	35
4	Rigorous power-law bounds on coarsening rates — The Kohn-Otto method	40
4.1	Basic inequalities	41
4.2	Bounds on coarsening rates for the LSW mean-field model	43
4.3	Bounds on coarsening rates for the monopole model	44
5	Smoluchowski's coagulation equations	47
5.1	Introduction	47
5.2	A 'new' framework for dynamic scaling analysis	47
5.3	Solution by Laplace transform	49
5.4	Scaling solutions and domains of attraction	51
5.5	The scaling attractor	54
5.6	Linearization of dynamics on the scaling attractor	57
	References	59

## 1. Introduction

**1.1.** These lectures are intended to serve as a rough introduction to a variety of mathematical problems aimed at understanding dynamic behavior in certain complex nonlinear physical systems, systems described by an evolution equation

$$\partial_t u = A(u), \quad (1.1)$$

where  $u(t)$  is the state of the system at time  $t$ , with  $u(t) \in X$ , a high-dimensional state space. In general, one may think of problems to do with statistical mechanics ( $\dim X \sim 10^{24}$ ), fluid turbulence, polycrystalline grain structure of typical metals, etc. There are many ways to approach the question of dynamics in such systems—experimental, computational, statistical, etc. We focus on the role of mathematical analysis, whose appropriate uses include: identifying general principles for dynamic behavior; detailed study of prototypical examples and critical cases; developing and justifying procedures to reduce complex systems to simpler ones (of lower dimension, perhaps); and identifying and studying significant structure (Hamiltonian dynamics, gradient flow, thermodynamic compatibility).

From the physical point of view, our particular focus is on models of kinetic behavior for systems whose spatial structure develops a pattern of domains or clusters that coarsen as time increases, in ways that seem to be *statistically self-similar*. This kind of behavior is seen, for example, in foams (bubble bath), grain structure in alloys, and many agglomeration and clustering processes. Although many of these systems are subject to the second law of thermodynamics—entropy increases (or at constant temperature, free energy decreases)—they do not reach equilibrium on the time scales of interest, and so it is an interesting problem to understand the regularities observed. If there is one universal theme of these lectures, it is that ultimately the universe may be doomed to heat death, but the path it takes along the way could be interesting nevertheless.

From the mathematical point of view, the emphasis is on studying dynamical phenomena peculiar to infinite-dimensional, spatially extended systems, such as weak convergence and self-similarity (dynamic scaling behavior). In infinite dimensions, the choice of an appropriate topology becomes nontrivial, and sometimes depends upon both mathematical and physical considerations.

**Note.** These lecture notes grew from tutorial lectures given in January 2005 in conjunction with the program on Nanoscale Material Interfaces

held at the Institute for Mathematical Sciences, University of Singapore, and were delivered in a course of lectures given at Humboldt University in Berlin in summer 2005. The notes were prepared with the assistance of Apostolos Damialis, Simone Hock, and Dirk Peschka. The lectures were developed for an audience including graduate and advanced undergraduate students; therefore a minimum of background in PDE theory is assumed, and the derivation of models and discussion of PDE results is largely done in a heuristic way. But a number of new rigorous results for models of the dynamics of domain size distributions are presented. Gradient structure turns out to play a surprisingly interesting and significant role throughout.

## 1.2. First models of domain formation and an open problem.

(i) A scalar ODE  $\partial_t u = -f(u)$  with  $f : \mathbb{R} \rightarrow \mathbb{R}$   $C^1$  always gives a gradient flow:  $\partial_t u = -W'(u)$  where  $W(u) = \int f(u) du$ . Every solution  $t \mapsto u(t)$  is monotonic, and every bounded solution converges as  $t \rightarrow \infty$ .

(ii) We introduce spatial variation, considering solutions of

$$\partial_t u(x, t) = -f(u(x, t)), \quad x \in \Omega = [0, 1], \quad t \geq 0. \quad (1.2)$$

What happens as  $t \rightarrow \infty$  is simple enough to describe: For any bounded solution,  $u_\infty(x) = \lim_{t \rightarrow \infty} u(x, t)$  exists for every  $x$ , with  $f(u_\infty(x)) = 0$  for all  $x$ . If  $f$  has multiple stable zeros (for example,  $f(u) = u^3 - u$ ), the limiting state  $u_\infty$  is typically non-constant—domains will form in  $\Omega$  as time proceeds, corresponding to different limiting values of  $u_\infty(x)$ .

Even for this simple equation, it is problematic to choose an appropriate space and topology for studying these infinite-dimensional dynamics. For state space we could try  $X = C(\Omega)$  with norm  $\|u\|_X = \sup_x |u(x)|$ . This is fine for proving local solvability of the initial value problem, but is a poor choice for studying long-time behavior. The pointwise limit  $u_\infty$  may be discontinuous, thus not in  $X$ . One can achieve  $u_\infty \in X$  by taking  $X = B(\Omega)$ , the space of bounded functions on  $\Omega$  with the sup norm. But then it is false in general that  $\|u(\cdot, t) - u_\infty\|_X \rightarrow 0$  as  $t$  gets large. One has  $\int_\Omega |u(\cdot, t) - u_\infty|^p \rightarrow 0$  for any finite  $p \geq 1$ , which suggests taking  $X = L^p(\Omega)$ . But this is not wonderful either—the Nemytskii operator  $f : L^p(\Omega) \rightarrow L^p(\Omega)$  is not usually  $C^1$ , and local solvability for all initial data may require restrictive hypotheses on  $f$  (globally Lipschitz, say).

(iii) A small change in the problem leads us quickly to the realm of the unknown. We alter the previous model by introducing a scalar “mean-field” coupling parameter  $\theta(t)$ , determined by imposing the global constraint that total “mass”  $\int_\Omega u(x, t) dx$  should remain conserved in time. Thus we con-

sider

$$\partial_t u(x, t) = -f(u(x, t)) + \theta(t), \quad x \in \Omega = [0, 1], \quad t \geq 0.$$

Mass conservation formally implies  $\theta(t) = \int_{\Omega} f(u(x, t)) \, dx$ . Assuming  $f : \mathbb{R} \rightarrow \mathbb{R}$  is  $C^1$ , the map  $u \mapsto \int_{\Omega} f(u)$  is  $C^1$  on the state space  $X = B(\Omega)$ , and Picard iteration gives us local-in-time solvability for initial-value problems

$$\partial_t u = -f(u) + \int_{\Omega} f(u), \quad u(t_0) = u_0 \in X.$$

One can rather easily prove global existence forward in time for bounded solutions under the mild hypothesis that

$$\liminf_{z \rightarrow -\infty} f(z) < f(s) < \limsup_{z \rightarrow +\infty} f(z) \quad \text{for every } s \in \mathbb{R}.$$

This leads us directly to the following innocent-sounding

**Open question:** Under only these mild hypotheses on  $f$ , must every bounded solution converge (pointwise a.e.) as  $t \rightarrow \infty$ ?

Various partial results are known from studying related PDEs from viscoelasticity [4] and anomalous diffusion [43]. But an answer to the general question remains elusive. Tantalizing is the fact that “free energy” is decreasing:

$$\frac{d}{dt} \int_{\Omega} W(u) = \int_{\Omega} W'(u)(\partial_t u) = \int_{\Omega} \left( f(u) - \int_{\Omega} f(u) \right) \partial_t u = - \int_{\Omega} (\partial_t u)^2,$$

thus

$$\int_{\Omega} W(u(t)) + \int_{t_0}^t \int_{\Omega} (\partial_t u)^2 = \int_{\Omega} W(u(t_0)).$$

Then it follows that as  $t$  increases,  $\int_{\Omega} W(u(t))$  decreases and has a limit, and one can show  $\int_{\Omega} (\partial_t u)^2$  is Lipschitz in  $t$  and tends to zero as  $t \rightarrow \infty$ . This rules out many types of recurrent dynamics such as asymptotically periodic behavior, but ever-slower drift remains a mathematical possibility. The main difficulty appears to be that it is not known whether  $\lim_{t \rightarrow \infty} \int_{\Omega} f(u(x, t)) \, dx$  must exist.

In [47] it is proved that *if the initial data  $u_0$  has finite range*, then indeed  $\lim_{t \rightarrow \infty} u(x, t)$  exists for all  $x \in \Omega$ . In this case, the problem reduces to one for a finite-dimensional ODE system for the unknown  $\vec{u}(t) = (u_j(t)) \in \mathbb{R}^N$ , where

$$u_j(t) = u(x, t) \quad \text{for } x \in A_j \subset \Omega, \quad \cup_{j=1}^N A_j = \Omega.$$

The proof exploits this finite-dimensionality in an essential way. The main ideas go as follows (see [47] for details): The  $\omega$ -limit set of the solution trajectory,

$$\omega(\bar{u}) = \bigcap_{t_0 > 0} \overline{\{\bar{u}(t) \mid t \geq t_0\}} = \{\bar{v} \in \mathbb{R}^N : \exists t_n \rightarrow \infty, \bar{u}(t_n) \rightarrow \bar{v}\},$$

is a connected, compact set. By Sard's theorem,  $f$  has an open, dense set of regular values. If  $\omega(\bar{u})$  is not a single point, then one can show that it contains a hyperbolic curve of equilibria. Using a theorem of Hale and Massat [25] related to center manifold theory, it follows that  $\omega(\bar{u})$  must contain a non-equilibrium point. But for this system,  $\omega(\bar{u})$  contains only equilibria, giving a contradiction.

## 2. A hierarchy of domain coarsening models in one space dimension

### 2.1. Domain walls in the Allen-Cahn equation

In a variety of physical processes, domains that form in multi-stable systems slowly change in time, with the overall pattern becoming coarser. Important examples in materials science include the growth of single-crystal grains in polycrystalline materials, phase separation in alloys, and anti-phase boundary motion in antiferromagnetic materials. One of the simplest mathematical models of this behavior arises as a modification of (1.2) above. Namely we consider the Allen-Cahn equation (or scalar Ginzburg-Landau equation)

$$\partial_t u(x, t) = -f(u) + \varepsilon^2 \partial_x^2 u, \quad 0 < x < 1, \quad t > 0. \quad (2.1)$$

The term  $\varepsilon^2 \partial_x^2 u$  can arise from diffusion, or through continuum modeling of nearest-neighbor coupling effects in atomic lattices, for example [3]. We have in mind that  $\varepsilon$  could be quite small, representing a ratio of microscopic to macroscopic length scales, for example. We fix attention on the bi-stable nonlinearity

$$f(u) = W'(u), \quad \text{where} \quad W(u) = \frac{1}{8}(u-1)^2(u+1)^2, \quad (2.2)$$

and for convenience impose the Neumann boundary conditions

$$\partial_x u = 0 \quad \text{at } x = 0 \text{ and } 1, \quad t > 0. \quad (2.3)$$

Our aim in this section is to describe the process of domain wall formation that occurs for this system, and how coarsening by domain wall motion and annihilation can be described via a hierarchy of models at varying levels of description, leading to a *universal kind of statistical behavior* on very large scales.

From the PDE-theoretic point of view it is convenient to take the state space  $X$  as the Sobolev space  $H^1([0, 1])$  of functions on  $[0, 1]$  with square-integrable first derivatives, but for these notes, we ignore PDE technicalities. For the most part, one may suppose solutions are smooth and take  $X = C^\infty([0, 1])$ .

For smooth solutions of the Allen-Cahn equation, the “free energy”

$$\mathcal{F}(u) = \int_0^1 \left( W(u(x)) + \frac{\varepsilon^2}{2} (\partial_x u)^2 \right) dx \quad (2.4)$$

decreases in time:

$$\frac{d}{dt} \mathcal{F}(u(t)) = \int_0^1 W(u) \partial_t u + \varepsilon^2 \partial_x u \partial_{tx} u = - \int_0^1 (\partial_t u)^2.$$

A fact of significance to us is that the Allen-Cahn equation has the structure of gradient flow for this free energy, with respect to the  $L^2$  inner-product on  $[0, 1]$  given by  $\langle u, v \rangle_{L^2} = \int_0^1 u(x)v(x) dx$ . Namely, if  $u, v \in X$  and  $u$  satisfies the boundary conditions, then

$$d\mathcal{F}(u)v = \frac{d}{d\tau} \mathcal{F}(u + \tau v)|_{\tau=0} = \int_0^1 (W'(u)v + \varepsilon^2 \partial_x u \partial_x v) dx = \langle \nabla_u \mathcal{F}, v \rangle_{L^2},$$

where the formal gradient  $\nabla_u \mathcal{F}(u) = W'(u) - \varepsilon^2 \partial_x^2 u$ . So formally,

$$\partial_t u = -\nabla_u \mathcal{F}(u). \quad (2.5)$$

Theory developed in the 1980's established some satisfying facts regarding the behavior of solutions in the long-time limit  $t \rightarrow \infty$ : For every solution,  $u_\infty = \lim_{t \rightarrow \infty} u(t)$  exists and satisfies the equation of equilibrium:

$$W'(u) - \varepsilon^2 \partial_x^2 u = 0. \quad (2.6)$$

There are exactly two stable equilibrium states, and no more:  $u \equiv 1$  and  $u \equiv -1$ . Unstable equilibria are restrictions of periodic solutions of the equation of equilibrium, which has first integral  $\varepsilon^2 v^2/2 - W(u) = W_0$ ,  $v = \partial_x u$ .

**Domain wall formation.** These results contrast sharply with one's expectations based on the model with  $\varepsilon = 0$  in (ii) above, and with numerical computations performed with small  $\varepsilon > 0$ . One expects and computes that  $u$  approaches 1 where  $u > 0$  initially, and  $u$  approaches  $-1$  where  $u < 0$  initially. “Domain walls” or transition layers form between these domains, at positions corresponding roughly to zeros in the initial data (so not necessarily periodically arranged).

**Structure of a domain wall.** The structure of a domain wall, located near position  $h \in (0, 1)$ , has a characteristic width of order  $\varepsilon$ , and can be

described crudely in terms of the stretched variable  $y = (x - h)/\varepsilon$  as a solution  $u = \pm\Theta(y)$  of the rescaled equilibrium equation

$$W'(\Theta) - \partial_y^2 \Theta = 0, \quad -\infty < y < \infty, \quad \Theta(y) \rightarrow \pm 1 \text{ as } y \rightarrow \pm\infty.$$

Explicitly one finds  $\Theta(y) = \tanh(y/2)$ . Note that as  $x \rightarrow \infty$ ,

$$\tanh(x/2\varepsilon) = 1 - 2\exp(-x/\varepsilon) + O(\exp(-2x/\varepsilon)).$$

Thus the domain structure that one expects to develop consists of arbitrarily placed domain walls of characteristic width  $\varepsilon$ , separating domains in which  $u$  is exponentially close to the stable states  $\pm 1$ . (For recent rigorous results on the various stages of domain wall formation and evolution, see [13].)

## 2.2. Domain wall dynamics by restricted gradient flow

Numerics indicates that the kind of domain wall pattern that develops as just described essentially stops<sup>a</sup> evolving, contrary to what theory says should happen as  $t \rightarrow \infty$ . What actually happens is that the domain walls typically move, albeit *extremely slowly*. A delicate formal analysis of the domain wall dynamics was made by J. Neu (unpublished notes), and this prompted subsequent development of rigorous geometric theory by Fusco and Hale [21] and in [11, 9].

The result, which we shall derive here formally using a restricted gradient flow approach, is that given  $N$  domain walls initially located at given positions  $h_1 < h_2 < \dots < h_N$  in  $(0, 1)$ , the positions will evolve in time according to equations well-approximated by exponentially small nearest-neighbor interactions:

$$\partial_t h_j = 12\varepsilon \left( \exp\left(-\frac{h_{j+1} - h_j}{\varepsilon}\right) - \exp\left(-\frac{h_j - h_{j-1}}{\varepsilon}\right) \right). \quad (2.7)$$

Here the wall positions  $h_0 = -h_1$  and  $h_{N+1} - 1 = 1 - h_N$  are obtained by reflection through the boundaries.

**Geometric description.** The Fusco-Hale idea for a geometric description of these slow dynamics is to describe solutions containing  $N$  domain walls in terms of an  $N$ -dimensional manifold of “metastable” states in  $X$ . We will parametrize a state  $u^h \in X$  on this manifold by a vector

<sup>a</sup>This depends upon the fact that the relative minima of  $W$  are equal. Different relative minima produce domain walls that move at speeds of order  $\varepsilon$  times the difference in minima.

$h = (h_1, \dots, h_N) \in \Delta_N$  of arbitrarily-placed well-separated domain wall positions. Here we set

$$\Delta_N = \{h \in \mathbb{R}^N \mid 0 < h_1 < \dots < h_N < 1, (h_{j+1} - h_j)/\varepsilon > K\},$$

where  $K$  is a large constant. Then

$$\mathcal{M}_N = \{u^h \mid h \in \Delta_N\} \subset X$$

is an  $N$ -dimensional manifold which should approximate the metastable states. Slowly evolving solutions will have a representation

$$u(x, t) = u^{h(t)}(x) + v(x, t) \tag{2.8}$$

where the error  $v(x, t)$  is ideally zero or approaching zero if  $\mathcal{M}$  is invariant.

**Restricted gradient flow.** Our formal approach here will be to construct  $\mathcal{M}_N$  as “approximately invariant” and compute equations of motion for the domain wall positions by restricting the gradient flow from (2.5) to the manifold  $\mathcal{M}_N$ .

Geometrically this can be interpreted in a couple of equivalent ways. One way is to simply project the right-hand side of the Allen-Cahn PDE (2.1) onto the tangent space  $T\mathcal{M}_N$ , using the  $L^2$  inner product on  $[0, 1]$ :  $\langle u, v \rangle_{L^2} = \int_0^1 u(x)v(x) dx$ . Thus we require

$$\langle \partial_t(u^h), v \rangle_{L^2} = \langle -\nabla_u \mathcal{F}(u^h), v \rangle_{L^2} \quad \text{for all } v \in T\mathcal{M}_N. \tag{2.9}$$

Since  $\partial_t(u^h) = \sum_{j=1}^N (\partial u^h / \partial h_j) \partial_t h_j$ , we get  $N$  equations for  $N$  unknowns  $\partial_t h_j$ .

An equivalent, more geometric, interpretation is useful to describe, directly in terms of a gradient flow

$$\partial_t h = -\nabla_h \tilde{\mathcal{F}}(h), \quad \text{where } \tilde{\mathcal{F}}(h) := \mathcal{F}(u^h). \tag{2.10}$$

This is a gradient flow on  $\Delta_N \subset \mathbb{R}^N$ , and the gradient is computed using the appropriate Riemannian metric pulled back from the tangent space  $T\mathcal{M}_N$ . Namely, given two tangent vectors  $a, \tilde{a} \in \mathbb{R}^N = T\Delta_N$ , we let

$$v = \sum_j a_j \frac{\partial u^h}{\partial h_j}, \quad \tilde{v} = \sum_j \tilde{a}_j \frac{\partial u^h}{\partial h_j}, \tag{2.11}$$

and define a metric  $g_h : T\Delta_N \times T\Delta_N \rightarrow \mathbb{R}$  by

$$g_h(a, \tilde{a}) = \langle v, \tilde{v} \rangle_{L^2} = \sum_{i,k} g_{ik} a_i \tilde{a}_k, \quad g_{ik} = \left\langle \frac{\partial u^h}{\partial h_i}, \frac{\partial u^h}{\partial h_k} \right\rangle_{L^2}.$$

The gradient  $\nabla_h \tilde{\mathcal{F}}(h)$  is determined from

$$g_h(\nabla_h \tilde{\mathcal{F}}(h), a) = d\tilde{\mathcal{F}}(h)(a) = \frac{d}{d\tau} \mathcal{F}(u^{h+\tau a})|_{\tau=0} = \sum_{k=1}^N \left\langle \nabla_u \mathcal{F}(u^h), \frac{\partial u^h}{\partial h_k} \right\rangle_{L^2} a_k.$$

The equations (2.9) and (2.10) are easily seen equivalent using the correspondence (2.11):

$$g_h(\partial_t h, a) = g_h(-\nabla_h \tilde{\mathcal{F}}(h), a) = \langle -\nabla_u \mathcal{F}(u^h), v \rangle_{L^2} = \langle \partial_t(u^h), v \rangle_{L^2}.$$

We can adequately approximate the metric coefficients  $g_{ik}$  formally using the domain-wall approximation that for  $x$  near  $h_j$ ,

$$u^h(x) \approx \pm \Theta \left( \frac{x - h_j}{\varepsilon} \right), \quad \frac{\partial u^h}{\partial h_j} \approx -\partial_x u^h \approx \mp \frac{1}{\varepsilon} \Theta' \left( \frac{x - h_j}{\varepsilon} \right).$$

Then

$$g_{ij} = \int_0^1 \frac{\partial u^h}{\partial h_i} \frac{\partial u^h}{\partial h_j} dx \approx 0 \quad (i \neq j), \tag{2.12}$$

$$g_{jj} = \int_0^1 \left( \frac{\partial u^h}{\partial h_j} \right)^2 dx \approx \frac{1}{\varepsilon} \int_{-\infty}^{+\infty} \Theta'(y)^2 dy \tag{2.13}$$

Since  $\frac{1}{2}\Theta_y^2 = W(\Theta) + 0$  one concludes

$$\int_{-\infty}^{+\infty} \Theta_y^2 dy = \int_{-1}^1 \sqrt{W(\theta)} d\theta = 2 \int_0^1 \frac{1 - \theta^2}{2} d\theta = \frac{2}{3}. \tag{2.14}$$

Thus the restricted metric is approximated by a diagonal matrix:

$$g_{ij} \approx \frac{2}{3\varepsilon} \delta_{ij}.$$

**Approximating manifold of metastable states.** The essential properties of the slowly evolving solutions are that  $W'(u) - \varepsilon^2 \partial_x^2 u \approx 0$  away from  $x = h_j$  and  $u(h_j, x) \approx 0$  for  $x \approx h_j$ . Thus we define  $\mathcal{M}_N$  as follows: Given  $h = (h_1, \dots, h_N) \in \Delta_N$  require for  $j = 0, \dots, N$  that

$$W'(u^h) - \varepsilon^2 \partial_x^2 u^h = 0 \text{ for } x \in I_j := (h_j, h_{j+1}), \tag{2.15}$$

$$u^h(x) = 0 \text{ for } x = h_j \text{ and } h_{j+1}, \tag{2.16}$$

$$(-1)^j u^h(x) > 0 \text{ in } I_j. \tag{2.17}$$

That is, in each interval  $I_j$  we have a piece of a periodic equilibrium with period  $2l_j$ , where  $l_j = (h_{j+1} - h_j)/\varepsilon$  measures domain length: Requiring

$\mathcal{U}(y, l)$  to satisfy

$$W'(\mathcal{U}) - \partial_y^2 \mathcal{U}(y, l) = 0, \tag{2.18}$$

$$\mathcal{U}(0, l) = \mathcal{U}(l, l) = 0, \tag{2.19}$$

$$\mathcal{U}(y, l) < 0 \quad \text{for } 0 < y < l, \tag{2.20}$$

we have

$$u^h(x) = (-1)^j \mathcal{U} \left( \frac{x - h_j}{\varepsilon}, \frac{h_{j+1} - h_j}{\varepsilon} \right).$$

Thus described, the states  $u^h$  are not smooth. They are continuous but  $\partial_x u^h$  is generally discontinuous at  $h_j$ . For this reason, the states  $u^h$  in [11, 9] were smoothed, trading one nuisance for another. (Results obtained without smoothing have been reported by Reznikoff (in preparation).)

Now, the restricted free energy turns out to be a sum of terms that depend only on the domain lengths: With

$$V(l) = \int_0^l \frac{1}{2} (\partial_y \mathcal{U})^2 + W(\mathcal{U}(y, l)) \, dy, \tag{2.21}$$

we can write

$$\tilde{\mathcal{F}}(h) = \mathcal{F}(u^h) = \sum_j \int_{I_j} \left( \frac{\varepsilon^2}{2} (\partial_x u^h)^2 + W(u^h) \right) dy = \sum_j \varepsilon V \left( \frac{h_{j+1} - h_j}{\varepsilon} \right).$$

We compute the gradient  $\nabla_h \tilde{\mathcal{F}}(h)$  as follows: Given  $a \in \mathbb{R}^N$ ,

$$\begin{aligned} \frac{d}{d\tau} \tilde{\mathcal{F}}(h + \tau a)|_{\tau=0} &= \sum_k V' \left( \frac{h_{k+1} - h_k}{\varepsilon} \right) (a_{k+1} - a_k) \\ &= \sum_k \left( V' \left( \frac{h_k - h_{k-1}}{\varepsilon} \right) - V' \left( \frac{h_{k+1} - h_k}{\varepsilon} \right) \right) a_k \\ &= \sum_k \frac{\partial \tilde{\mathcal{F}}}{\partial h_k} a_k = \sum_{j,k} g_{jk} \nabla_h \tilde{\mathcal{F}}(h)_j a_k. \end{aligned}$$

Using the diagonal approximation of the metric from above gives the approximate equations of motion

$$\partial_t h_j = -\nabla_h \tilde{\mathcal{F}}(h)_j \approx \frac{3\varepsilon}{2} \left( V' \left( \frac{h_{k+1} - h_k}{\varepsilon} \right) - V' \left( \frac{h_k - h_{k-1}}{\varepsilon} \right) \right) \tag{2.22}$$

The interpretation is that each domain wall moves as if attracted by each of its two nearest neighbors, with “force” determined by the potential  $V$ .

**Approximation of the force.** With  $V(l)$  given by (2.21), one obtains

$$\begin{aligned} V'(l) &= \left( \frac{1}{2}(\partial_y \mathcal{U})^2 + W(\mathcal{U}) \right)_{y=l} + \int_0^l ((\partial_y \mathcal{U})(\partial_{y^2}^2 \mathcal{U}) + W'(\mathcal{U})\partial_l \mathcal{U}) \, dy \\ &= \left( \frac{1}{2}(\partial_y \mathcal{U})^2 + W(\mathcal{U}) \right)_{y=l} + (\partial_y \mathcal{U})(\partial_l \mathcal{U}) \Big|_0^l \end{aligned}$$

after an integration by parts using (2.18). Recall that  $U(l, l) = 0$ , hence we have  $(\partial_y \mathcal{U} + \partial_l \mathcal{U})(l, l) = 0$ . Also  $\partial_l \mathcal{U}(0, l) = 0$ . Thus

$$V'(l) = \overbrace{\left( -\frac{1}{2}(\partial_y \mathcal{U})^2 + W(\mathcal{U}) \right)}^{\text{const. on } [0, l]} \Big|_{y=l} = W \left( \mathcal{U} \left( \frac{l}{2}, l \right) \right).$$

To approximate this, we use that for large  $l$ , we have  $\mathcal{U}(y + l/2, l) \approx -1$  when  $|y| \ll l/2$  (meaning  $0 \ll y + l/2 \ll l$ ), whence

$$\mathcal{U}(y + \frac{l}{2}, l) \approx -1 + \tilde{\mathcal{U}}(y) \quad \text{where} \quad \partial_y^2 \tilde{\mathcal{U}} - W''(-1)\tilde{\mathcal{U}} = 0.$$

Since  $W''(-1) = 1$  and  $\partial_y \mathcal{U}(l/2, l) = 0$ , we have  $\tilde{\mathcal{U}}(y) = \alpha(e^y + e^{-y})$ . On the other hand, near the right endpoint of  $[0, l]$ ,  $\mathcal{U}$  is approximated by the domain wall structure:

$$\mathcal{U}(\tilde{y} + l, l) \approx \Theta(\tilde{y}) = \tanh(\tilde{y}/2) \approx -1 + 2e^{\tilde{y}} \quad \text{for } l/2 \ll \tilde{y} + l \ll l.$$

To identify  $\alpha = \alpha(l)$ , we match these approximations in the regime  $0 \ll y = \tilde{y} + l/2 \ll l/2$ . This means we require  $\alpha e^{\tilde{y}+l/2} = 2e^{\tilde{y}}$ , whence  $\alpha = 2e^{-l/2}$  and consequently for large  $l$  we get the approximation

$$\begin{aligned} V'(l) &= W(\mathcal{U}(l/2, l)) \approx W(-1) + W'(-1)\tilde{\mathcal{U}} + \frac{1}{2}W''(-1)\tilde{\mathcal{U}}^2 \\ &= \frac{(2\alpha)^2}{2} = 8e^{-l}. \end{aligned}$$

Together with (2.22) this leads to the equations of motion in (2.7).

### 2.3. Punctuated equilibrium and 1D bubble bath

**Lifetime of metastable states.** Let us imagine a system in which  $1/\varepsilon$ , the ratio of macroscopic domain size to microscopic domain wall thickness, is quite large (say  $> 10^5$ ), and suppose that at some point the system settles into a metastable state with a great number of domain walls. Then due to the exponential dependence of terms in (2.7) on domain size, one

can expect the dynamics to be dominated by the smallest domains. Say  $l = \min(h_j - h_{j-1})$ , then approximately

$$\partial_t l = -24\varepsilon e^{-l/\varepsilon}, \quad \text{or} \quad \partial_t e^{l/\varepsilon} = -24.$$

According to this equation, the domain size  $l(t)$  shrinks to zero in a finite time  $T$  determined from  $\exp(l_0/\varepsilon) = 1 - 24T$ . The following table gives some indication of how the lifetime  $T$  of the metastable state depends strongly on the minimum initial domain size,  $l_0/\varepsilon$  (measured in units of domain wall thickness):

$l_0/\varepsilon$	5	10	20	50	100
$T$	4	918	$2 \times 10^7$	$2 \times 10^{20}$	$10^{42}$

In geometric terms, we have described the nature of  $N$ -wall metastable states using an approximately invariant manifold  $\mathcal{M}_N$ . (Of course there are two of these, using  $-u^h$ .) Interestingly, however, there is indeed an invariant manifold close to  $\mathcal{M}_N$ , which nearby solutions approach at a uniform rate as long as domain walls remain well separated [9]. That metastable states should correspond to such an invariant manifold was conjectured by Fusco and Hale [21], who also suggested that this manifold is part of the global *unstable manifold* of the unstable  $N$ -wall equilibrium state with equal domain sizes  $h_{j+1} - h_j$ . Indeed it was established in [9] that this unstable manifold is given in terms of (2.8) (with smoothed  $u^h$ ) as a graph  $h \mapsto v$  globally over  $\Delta_N$  with exponentially small Lipschitz constant.

**Punctuated equilibrium.** The analysis so far indicates that the story of how gradient systems relax to equilibrium is not as simple as looking at stable steady states (here only  $u \equiv 1$  and  $u \equiv -1$ ) and finding the local rate of approach to these states (which here is  $O(1)$ ). Instead, dynamics in the simple PDE (2.1) can be expected to exhibit a cascading behavior reminiscent of the “punctuated equilibrium” description of species evolution advocated by Stephen Jay Gould. (Some rigorous results along these lines were established in [18].) We might expect a typical solution trajectory to behave as follows. Domain walls develop and the system approaches an  $N$ -dimensional metastable manifold of states with  $N$  domain walls positioned arbitrarily as determined by initial data. It flows very slowly along this invariant manifold until two domain walls come close together (or one approaches the boundary). The two walls (or the one and its reflection) rapidly annihilate each other on an order 1 time scale, and the solution then settles into approaching an  $N - 2$  (or  $N - 1$ ) dimensional metastable manifold. The slow motion grows dramatically slower as fewer walls remain,

with greater distance between them. Note that the solution may be “near equilibrium” in the sense of being nearly stationary, but need never be near an equilibrium *state*. In the limit  $t \rightarrow \infty$ , one can expect to approach one of the stable states  $u \equiv \pm 1$ —but it may not be practical to wait that long!

**1D bubble bath.** So far we have reduced the study of long-time behavior in the Allen-Cahn PDE to the motion of domain walls according to the ODEs in (2.7), and have further simplified by noticing that the largest term in these ODEs should strongly dominate and produce collapse of the smallest domain. This suggests an even simpler model that we can use to investigate the statistical behavior of the coarsening process.

Starting by partitioning the interval  $[0, 1]$ , placing domain walls randomly according to some scheme. Then coarsen this domain pattern according to the following recipe:

- (1) The smallest domain joins its two neighbors.
- (2) Nobody else moves.
- (3) Repeat.

This process is a kind of 1D model of bubble bath—the smallest bubble pops first, and the foam becomes coarser in time. This 1D model of coarsening is easy to simulate by computer with many thousands of domains. (See [10]. Many results for related models also exist in the physical literature; see [17] for a review.) Results show a remarkable thing: After scaling by mean size, the distribution of domain sizes develops toward a *universal self-similar form*. This raises the interesting question:

**Why?**

#### **2.4. Mean-field model of domain growth — The Gally-Mielke transform**

To try and understand this phenomenon of universal self-similarity, we formulate a model of this domain coalescence process that aims to describe how the domain size distribution evolves in time. The main idea is to develop a rate equation for the domain size distribution function, based upon the mean-field assumption that the sizes of domains undergoing coalescence are accurately characterized by the overall size distribution of all domains. For the model that results, a remarkable solution procedure was recently developed by Gally and Mielke [22], which we will use in this section to prove a theorem regarding universal approach to self-similar form.

**Derivation of the model.** We let  $x$  denote domain size, and at first take  $x$  discrete:  $x = j\Delta x$ ,  $j \in \mathbb{N}$ . We introduce a function to describe the domain size distribution by

$$f(x, t)\Delta x = \text{expected number of domains of size } x \\ \text{(normalized by initial total)}$$

Let  $\mathcal{L}(t)$  denote the size of the smallest domain remaining at time  $t$ ; thus  $f(x, t) = 0$  for  $x < \mathcal{L}(t)$ . And let  $N(t) = \sum_x f(x, t)\Delta x$  be the total number at time  $t$ .

In the time interval  $(t, t + \Delta t)$ , the total number of coalescence events (involving the smallest remaining domain combining with its two neighbors) is expected to be

$$f(\mathcal{L}(t), t)\Delta \mathcal{L} = f(\mathcal{L}(t), t)\frac{\Delta \mathcal{L}}{\Delta t} \Delta t.$$

The change in number of size- $x$  domains will equal the total number of coalescence events times the sum over subevents of the relative probability of the subevent times the change in number of size- $x$  domains in the subevent. Three types of subevents affect size- $x$  domains:

- (1) Sizes  $(x, \mathcal{L}, y)$  combine to form  $x + \mathcal{L} + y$ .
- (2) Sizes  $(y, \mathcal{L}, x)$  combine to form  $y + \mathcal{L} + x$ .
- (3) Sizes  $(y, \mathcal{L}, x - y - \mathcal{L})$  combine to form  $x$ .

Under the mean field assumption, these events respectively have the relative probabilities

$$\frac{f(x)\Delta x}{N} \frac{f(y)\Delta x}{N}, \quad \frac{f(y)\Delta x}{N} \frac{f(x)\Delta x}{N}, \quad \frac{f(y)\Delta x}{N} \frac{f(x - y - \mathcal{L})\Delta x}{N}.$$

From these ideas, one finds the change in number of size- $x$  domains is

$$\Delta(f(x, y)\Delta x) = \\ f(\mathcal{L}, t)\Delta \mathcal{L} \sum_y \left( \frac{f(y, t)\Delta x}{N} \frac{f(x - y - \mathcal{L}, t)\Delta x}{N} - 2 \frac{f(x)\Delta x}{N} \frac{f(y)\Delta x}{N} \right)$$

Dividing by  $\Delta x \Delta t$  and passing formally to the continuum limit, one obtains

$$\partial_t f(x, t) = \frac{f(\mathcal{L}, t)\dot{\mathcal{L}}}{N^2} \int_0^\infty \left( f(y, t)f(x - y - \mathcal{L}, t) - 2f(x, t)f(y, t) \right) dy \quad (2.23)$$

This is the model rate equation we seek [36, 10].

Proceeding formally, we derive a useful moment identity and reformulate the rate equation. In the discrete case, we can write a general moment identity as follows:

$$\begin{aligned} \Delta \left( \sum_x a(x) f(x, t) \Delta x \right) &= -a(\mathcal{L}) f(\mathcal{L}, t) \Delta \mathcal{L} \\ &+ \frac{f(\mathcal{L}, t) \Delta \mathcal{L}}{N^2} \sum \left( a(\tilde{x} + y + \mathcal{L}) f(y) f(\tilde{x}) - 2 a(x) f(x) f(y) \right) \Delta x \Delta y \\ &= \frac{f(\mathcal{L}, t) \Delta \mathcal{L}}{N^2} \sum \left( a(x + y + \mathcal{L}) - a(x) - a(y) - a(\mathcal{L}) \right) f(x) f(y) \Delta x \Delta y. \end{aligned}$$

In the continuum limit this yields

$$\begin{aligned} \partial_t \int_{\mathcal{L}} a(x) f(x, t) dx & \tag{2.24} \\ &= \frac{f(\mathcal{L}, t) \dot{\mathcal{L}}}{N(t)^2} \int_0^\infty \int_0^\infty \left( a(x + y + \mathcal{L}) - a(x) - a(y) - a(\mathcal{L}) \right) f(x, t) f(y, t) dx dy \end{aligned}$$

Considering  $a(x) = x$  yields

$$\partial_t \int_0^\infty x f(x, t) dx = 0,$$

thus total size is preserved (if finite). Taking  $a(x) = 1$  next we get

$$\partial_t N(t) = -2 f(\mathcal{L}, t) \dot{\mathcal{L}}.$$

Thus total number decreases, and average domain size  $\bar{x} = \int x f / \int f$  increases. Also  $a(x) = x^2$  yields a growth law for second moment:

$$\partial_t \int_0^\infty x^2 f(x, t) dx = 2 f(\mathcal{L}, t) \dot{\mathcal{L}} (\bar{x}^2 + 2 \bar{x} \mathcal{L}).$$

Our model is invariant under reparametrization in time: If one changes variables via  $t = T(\tilde{t})$ ,  $\tilde{f}(x, \tilde{t}) = f(x, t)$ ,  $\tilde{\mathcal{L}}(\tilde{t}) = \mathcal{L}(t)$ , then the equation retains its form since

$$\partial_{\tilde{t}} \tilde{f} = \dot{T} \partial_t f, \quad \partial_{\tilde{t}} \tilde{\mathcal{L}} = \dot{T} \partial_t \mathcal{L}.$$

The model has no intrinsic time scale since the process is simply driven by the rate of collapse of smallest domains.

Following Gally and Mielke, it is convenient to parametrize time by the size of smallest domain, and take  $\mathcal{L}(t) = t$ . Also it is convenient to rewrite the model in terms of the probability density for domain size:

$$\rho(x, t) = \frac{f(x, t)}{N(t)}.$$

Since for  $\mathcal{L} = t$  we have

$$\partial_t \left( \frac{f(x, t)}{N(t)} \right) = \frac{\partial_t f}{N} - \frac{f \partial_t N}{N^2} = \frac{\partial_t f(x, t)}{N} + \frac{f(x, t)}{N} \cdot \frac{2f(t, t)}{N},$$

the model takes the form

$$\partial_t \rho(x, y) = \rho(t, t) \int_0^{x-t} \rho(y, t) \rho(x - y - t, t) dy \quad \text{for } x > t, \quad (2.25)$$

with  $\rho(x, t) = 0$  for  $x < t$ . Note that due to the latter condition, the integrand vanishes unless  $t < y < x - 2t$ , requiring  $x > 3t$ .

**The Gally-Mielke global linearizing transform.** An amazing solution procedure for this model was found by Gally and Mielke in [22], and used to establish several results regarding convergence to self-similar form at various rates depending upon the tail of the initial data. In these notes our aim is to give a simple proof of universal weak convergence to self-similar form for all classical solutions with finite total number and size.

Consider the initial value problem for the model (2.25), with initial data given at a time when smallest domain size  $t = 1$ , say. Leaving aside til later the question of solvability of this initial value problem, let us describe the solution procedure of Gally and Mielke. For brevity we use the (probabilists') notation  $\rho_t(x) = \rho(x, t)$  to denote the solution at time  $t$ , and denote its Fourier transform by

$$\mathcal{F}\rho_t = \hat{\rho}(\xi, t) = \int_{\mathbb{R}} e^{-i\xi x} \rho(x, t) dx.$$

With  $\Phi(z) = \frac{1}{2} \ln \frac{1+z}{1-z} = \tanh^{-1} z$ , introduce the change of variables

$$v_t(x) = \mathcal{F}^{-1} \circ \Phi \circ \mathcal{F}\rho_t, \quad \text{so} \quad \hat{v}_t(\xi) = \frac{1}{2} \ln \left( \frac{1 + \hat{\rho}_t}{1 - \hat{\rho}_t} \right).$$

In terms of  $v_t$  the solution is given by the simple formula (!)

$$v_t(x) = H(x - t)v_1(x) = \begin{cases} v_1(x), & x \geq t, \\ 0, & x \leq t, \end{cases}$$

where  $H$  is the Heaviside function. Thus, to find the solution to the non-linear model (2.25) at time  $t$ , the procedure is:

- (i) Transform the initial data  $\rho_1(x) = \rho(x, 1)$ : Let  $v_1(x) = \mathcal{F}^{-1} \circ \Phi \circ \mathcal{F}\rho_1$ .
- (ii) Set to zero for  $x \leq t$ : Let  $v_t(x) = H(x - t)v_1(x)$ .
- (iii) Invert the transformation:  $\rho_t = \mathcal{F}^{-1} \circ \Phi^{-1} \circ \mathcal{F}v_t = \mathcal{F}^{-1}(\tanh \mathcal{F}v_t)$ .

Let us now formally derive this method. I prefer to work with the Laplace transform to simplify rigorous analysis later. Denote the Laplace transform of  $\rho_t$  by

$$R_t(q) = \mathcal{L}\rho_t(q) = \int_0^\infty e^{-qx} \rho_t(x) dx.$$

The integral term on the right-hand side of (2.25) is a shifted convolution, and its Laplace transform is given by

$$\begin{aligned} \mathcal{L}(\rho * \rho(\cdot - t)) &= \int_0^\infty \int_0^{x-t} e^{-q(x-t-y+t+y)} \rho_t(y) \rho_t(x-t-y) dy dx \\ &= \int_0^\infty \int_0^\infty e^{-q\tilde{x}} e^{-q\tilde{y}} e^{-qt} \rho_t(\tilde{y}) \rho_t(\tilde{x}) d\tilde{y} d\tilde{x} \\ &= e^{-qt} R_t(q)^2. \end{aligned}$$

Since  $\partial_t R_t = \partial_t \int_t^\infty e^{-qt} \rho(x, t) dx = -e^{-qt} \rho(t, t) + \int_0^\infty e^{-qx} \partial_t \rho dx$ , taking the Laplace transform of (2.25) yields

$$\partial_t R_t = \alpha(t) e^{-qt} (-1 + R_t^2), \quad \text{where } \alpha(t) = \rho(t, t). \tag{2.26}$$

This yields

$$\partial_t \Phi(R_t) = \frac{\partial_t R_t}{1 - R_t^2} = -\alpha(t) e^{-qt},$$

whence upon integration,

$$\Phi(R_t) - \Phi(R_1) = - \int_1^t \alpha(s) e^{-qs} ds. \tag{2.27}$$

Note that for  $q > 0$ ,

$$R_t(q) = \int_t^\infty e^{-qx} \rho_t(x) dx \leq e^{-qt} \int_t^\infty \rho_t(x) dx = e^{-qt} \cdot 1 \rightarrow 0 \text{ as } t \rightarrow \infty.$$

Hence taking  $t \rightarrow \infty$  in (2.27) above yields

$$\Phi(R_1) = \int_1^\infty \alpha(s) e^{-qs} ds = \mathcal{L}\alpha. \tag{2.28}$$

This formula determines  $\alpha = \rho(\cdot, \cdot)$  in terms of the initial data, according to

$$\alpha = \mathcal{L}^{-1} \circ \Phi \circ \mathcal{L} \circ \rho_1.$$

Plugging back into (2.27) yields

$$\Phi(R_t) = \int_t^\infty e^{-qs} \alpha(s) ds = \mathcal{L}(H(\cdot - t)\alpha) = \mathcal{L}(\alpha_t), \tag{2.29}$$

where we set  $\alpha_t(x) = H(x-t)\alpha(x)$  for all  $x$ . Since  $\tanh(\Phi(z)) = z$ , it follows

$$R_t = \tanh(\mathcal{L}\alpha_t), \tag{2.30}$$

that is,  $\rho_t = \mathcal{L}^{-1} \circ \tanh \circ \mathcal{L}\alpha_t$ . This finishes the derivation. (In fact,  $\alpha_t = v_t$ .)

**Initial value problem.** Here we sketch a proof of existence for classical solutions of the initial value problem for the model (2.25), and a rigorous justification of the solution formula (2.30). We fix  $\tau = 1$  and suppose that  $\rho_\tau : [\tau, \infty) \rightarrow \mathbb{R}$  is given as a continuous function with  $\int_\tau^\infty \rho_\tau(x) dx = 1$ .

Note that since the solution is to satisfy  $\rho_t(x) = 0$  for  $x < t$ , the convolution term on the right-hand side of (2.25) will depend only upon values of  $\rho_t(y)$  for  $\tau < y < x - 2t \leq x - 2\tau$ . In particular, the right-hand side vanishes for  $x < 3\tau$ .

This means we can construct the solution for  $\tau < t < 3\tau$  by an inductive procedure as follows: For  $\tau < t \leq x < 3\tau$  we have  $\rho_t(x) = \rho_\tau(x)$  and in particular  $\rho_t(t) = \rho_\tau(t)$ . For  $\tau < t \leq 3\tau$ , successively on strips  $x \in [k\tau, (k + 2)\tau]$ , for  $k = 3, 5, \dots$ , by simple integration in time we can now compute

$$\rho_t(x) = \rho_\tau(x) + \int_\tau^t \rho_s(s) \int_0^{x-s} \rho_s(y)\rho_s(x - y - s) dy ds,$$

where the right-hand side is always known from a previous step. This determines  $\rho_t(x)$  for  $\tau \leq t \leq 3\tau$  and all  $x$ .

To determine the solution globally for all  $t > 1$ , the idea is to replace  $3\tau$  by  $\tau$  and repeat. But in order to justify this we need to verify that  $\rho_t$  remains integrable and conserves total probability. In particular we need to justify (2.26). Let us introduce the distribution function

$$F_t(x) = \int_0^x \rho_t(y) dz. \tag{2.31}$$

This is the probability that a domain has size  $\leq x$  at time  $t$ . Integrating the convolution term in (2.25), we get

$$\begin{aligned} \int_0^x \int_0^\infty \rho_t(y)\rho_t(z - y - t) dy dz &= \int_{\mathbb{R}} F_t(x - y - t) F_t(dy) \\ &\leq F_t(x) \int_0^x F_t(dy) = F_t(x)^2 \end{aligned}$$

since  $x \mapsto F_t(x)$  is increasing. Thus  $\partial_t F_t(x) \leq \rho_t(t)(F_t(x)^2 - 1)$ , and since  $\rho_t(t) \geq 0$  and  $F_t(x) \leq 1$  initially,  $F_t(x)$  is decreasing in  $t$  for fixed  $x$ . It follows  $F_t(\infty) \leq 1$ , and so the Laplace-Stieltjes transform

$$R_t(q) = \int_0^\infty e^{-qx} F_t(dx) = \int_t^\infty e^{-qx} \rho_t(x) dx$$

is well defined and  $\leq e^{-qt}$ . Since  $\partial_t F_t(x)$  is continuous in  $t$  for all  $x$ ,  $R_t(q)$  is  $C^1$  in  $t$  for all  $q > 0$ . This justifies (2.26) and the solution formula (2.30) in the previous subsection.

**Background on Laplace transforms.** Our rigorous study of dynamic scaling behavior will make use of some basic facts regarding Laplace transforms of measures on  $[0, \infty)$ . For this material we refer to Feller's excellent book [19]. In particular we recall the following main results from chapters VIII and XIII of [19]:

1. (Selection theorem) Every sequence of probability distributions has a subsequence that converges (weakly, i.e., in distribution) to some limit distribution (possibly defective).
2. (Continuity theorem) Weak convergence of measures is equivalent to pointwise convergence of the corresponding Laplace transforms.
3. (Tauberian theorem) Let  $U$  be a measure on  $[0, \infty)$  with  $U(0) = 0$  and suppose its Laplace-Stieltjes transform is

$$\omega(q) = \int_0^\infty e^{-qx} U(dx).$$

Let  $p \in [0, \infty)$ , and let  $L$  be a function *slowly varying at  $\infty$* , meaning  $\lim_{t \rightarrow \infty} L(tx)/L(t) = 1$  for all  $x > 0$ . Then the following are equivalent:

- (i)  $U(t) = U([0, t]) \sim t^p L(t)$  as  $t \rightarrow \infty$ .
- (ii)  $\omega(q) \sim q^{-p} L(1/q) \Gamma(1+p)$  as  $q \rightarrow 0$ .

For later reference we also now recall a fundamental lemma on scaling limits:

4. (Rigidity of scaling limits) Suppose  $U : [0, \infty) \rightarrow \mathbb{R}$  is positive and increasing, and suppose the following limit exists:

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = \psi(x) \leq \infty,$$

for all  $x$  in some set  $S$  dense in  $[0, \infty)$ . Then necessarily the limit is a power law:  $\psi(x) = x^p$  for some  $p \in [0, \infty]$ , and furthermore,  $U$  is *regularly varying at  $\infty$  with exponent  $p$* , meaning  $U(t) = t^p L(t)$  where  $L$  is slowly varying at  $\infty$ .

## 2.5. Proof of universal self-similar behavior

The goal here is to prove that every solution of the coarsening model (2.25) with initially finite expected size  $\int_0^\infty x \rho_1(x) dx$  will converge in distribution to a universal self-similar form. One can think of this as a dynamic analog of

the *central limit theorem* in probability. See section 5.2 below for a further development of this analogy.

**Rescaling.** In studying dynamic scaling behavior for this model, it is natural to rescale to keep the smallest domain size fixed. Hence we introduce the rescaled probability distribution function

$$\eta_t(x) = F_t(tx) = \int_0^{tx} \rho_t(y) dy. \tag{2.32}$$

We have  $\eta_t(x) = 0$  for  $x < 1$ , and  $\eta_t(\infty) = 1$ . Its Laplace transform is

$$N_t(q) = \int_1^\infty e^{-qx} \eta_t(dx) = \int_t^\infty e^{-qx/t} F_t(dx) = \mathcal{L}F_t(q/t)$$

We introduce the notation

$$A(t) = \int_1^t \alpha(s) ds = \int_1^t \rho_s(s) ds, \quad \text{so} \quad A(ds) = \alpha(s) ds.$$

By the solution formula (2.30) we have that

$$N_t(q) = \tanh \left( \int_1^\infty e^{-qs} A(t ds) \right). \tag{2.33}$$

**Self-similar solutions.** For self-similarity,  $N_t \equiv N$  is independent of  $t$ . Since then we have  $H(s - t)A(t ds) = A(ds)$  for all  $t$ , we must have

$$A(ds) = \frac{\beta}{s} H(s - 1) ds$$

for some constant  $\beta > 0$ . This means that the Laplace transform of the profile

$$N(q) = \tanh \left( \beta \int_1^\infty \frac{e^{-qs}}{s} ds \right) = \tanh(\beta \text{Ei}(q)), \tag{2.34}$$

where  $\text{Ei}(q) = \int_q^\infty (e^{-s}/s) ds$  is the exponential integral function.

Only certain values of  $\beta$  are physically meaningful here. Note that:

(i) To have  $\bar{y} = \int_1^\infty y\eta(dy) < \infty$  we need

$$-\partial_q N(q) = \int_1^\infty e^{-qy} y \eta(dy) \longrightarrow \bar{y} < \infty \quad \text{as } q \rightarrow 0.$$

(ii)  $\frac{d}{dx} \tanh x = \text{sech}^2 x = \frac{4}{(e^x + e^{-x})^2} = \frac{4e^{-2x}}{(1 + e^{-2x})^2}$ .

(iii)  $\text{Ei}(q) = \int_q^\infty \frac{e^{-s}}{s} ds = -\ln q + \gamma(q)$ , where  $\gamma(0) \approx 0.577216$  is Euler's constant.

(iv)  $\exp(-\beta \text{Ei}(q)) = q^\beta e^{\beta\gamma(q)}$ , hence as  $q \rightarrow 0$ ,

$$-\partial_q N(q) = \text{sech}^2(\beta \text{Ei}(q)) \frac{\beta e^{-q}}{q} = \frac{4\beta e^{-q}}{q} \frac{q^{2\beta} e^{2\beta\gamma}}{(1 + q^{2\beta} e^{2\beta\gamma})^2} \sim cq^{2\beta-1}.$$

Thus, we find that it is *necessary* that

$$\begin{aligned} \beta &> 0, \text{ for } N(q) > 0, \\ \beta &\leq \frac{1}{2}, \text{ for } -\partial_q N(q) = \int_0^\infty e^{-qy} y \eta(dy) \text{ to decrease in } q, \\ \beta &= \frac{1}{2}, \text{ for } -\partial_q N(q) \rightarrow \bar{y} \in (0, \infty) \text{ as } q \rightarrow 0. \end{aligned}$$

So there is a *unique possibility* for a self-similar solution with finite expected domain size, namely with  $\beta = 1/2$  and

$$N(q) = \tanh\left(\frac{1}{2} \text{Ei}(q)\right). \tag{2.35}$$

**Main result.** At this point we need to address the following questions:

- Does a positive self-similar solution really exist satisfying (2.35)?
- Is it stable, and does it attract every solution?

The answers are positive.

**Theorem 2.1:** *Suppose that the initial data for model (2.25) satisfy  $\int_0^\infty x \rho_1(dx) < \infty$ . Then with  $\eta_t(x)$  given by (2.32), we have  $\lim_{t \rightarrow \infty} \eta_t(x) = \eta_*(x)$  for all  $x > 1$ , where  $\eta_*$  is a probability distribution function satisfying  $\mathcal{L}\eta_*(q) = \tanh(\frac{1}{2} \text{Ei}(q))$ .*

**Proof:** 1. By the selection and continuity theorems for Laplace transforms, it suffices to show  $N_t(q) \rightarrow \tanh(\frac{1}{2} \text{Ei}(q))$  for all  $q > 0$ , or equivalently by (2.33),

$$\int_1^\infty e^{-qs} A(t ds) \longrightarrow \int_1^\infty \frac{e^{-qs}}{2s} ds \quad \text{as } t \rightarrow \infty \text{ for all } q > 0. \tag{2.36}$$

2. Let  $\bar{x} = \int_1^\infty y \eta_1(dy)$ . Then as  $q \rightarrow 0$ ,  $-\partial_q N_1(q) = \int_1^\infty e^{-qx} x \rho_1(dx) \rightarrow \bar{x}$ , and thus  $N_1(q) = 1 - q\bar{x}(1 + \sigma(1))$ . We have  $\mathcal{L}A(q) = \int_1^\infty e^{-qs} A(ds) = \Phi(N_1(q))$  by (2.28), hence

$$\begin{aligned} -\partial_q \mathcal{L}A(q) &= \int_1^\infty e^{-qs} s A(ds) = \frac{-\partial_q N_1(q)}{(1 + N_1(q))(1 - N_1(q))} \\ &= \frac{\bar{x}}{(1 + 1)(q\bar{x})} \cdot (1 + \sigma(1)) = \frac{1}{2q}(1 + \sigma(1)) \end{aligned}$$

as  $q \rightarrow 0$ . By the Tauberian theorem, it follows that the distribution function for the measure  $sA(ds)$  satisfies

$$\int_0^t sA(ds) = \frac{t}{2}(1 + \sigma(1)) \quad \text{as } t \rightarrow \infty.$$

3. Thus the distribution function for the measure  $sA(t ds)$  satisfies

$$\mu_t(x) = \int_0^x sA(t ds) = \frac{1}{t} \int_0^{tx} sA(ds) \rightarrow \frac{x}{2} \quad \text{as } t \rightarrow \infty,$$

i.e.,  $\mu_t \rightarrow \frac{1}{2}$  in distribution. It remains to show that for all  $q > 0$ ,

$$\int_1^\infty \frac{e^{-qs}}{s} \mu_t(ds) \rightarrow \int_1^\infty \frac{e^{-qs}}{s} \frac{ds}{2} \quad \text{as } t \rightarrow \infty.$$

We establish this in two steps: (i) By the weak convergence theorem for probability measures, given  $\bar{x} > 0$ , since the probability distribution function

$$\nu_t(x) = \min(\mu_t(x)/\mu_t(\bar{x}), 1) \rightarrow \min(x/\bar{x}, 1) \quad \text{as } t \rightarrow \infty,$$

we have  $\int_0^\infty u(s)\mu_t(ds)/\mu_t(\bar{x}) \rightarrow \int_0^\infty u(s) ds/\bar{x}$  for all  $u \in \mathcal{C}_c([0, \bar{x}])$  (continuous  $u$  with compact support in  $[0, \bar{x}]$ ), and hence

$$\int_0^\infty u(s)\mu_t(ds) \rightarrow \int_0^\infty u(s) \frac{ds}{2}.$$

This holds also for discontinuous  $u$ , provided that  $\int_0^\infty u = \inf \int_0^\infty v_+ = \sup \int_0^\infty v_-$ , where the inf and sup are taken over all  $v_- \leq u \leq v_+$  with  $v_+, v_- \in \mathcal{C}_c([0, \bar{x}])$ . Hence for all  $x > 1$ , we can conclude that

$$\int_1^x \frac{e^{-qs}}{s} \mu_t(ds) \rightarrow \int_1^x \frac{e^{-qs}}{s} \frac{ds}{2}.$$

(ii) We compute that

$$\begin{aligned} \int_0^\infty e^{-qs} \mu_t(ds) &= \int_0^\infty e^{-qs} sA(t ds) = \frac{1}{t} \int_0^\infty e^{-qs/t} sA(ds) \\ &= \frac{1}{t} \frac{-\partial_q N_1(q/t)}{(1 + N_1)(1 - N_1(q/t))} \leq \frac{1}{t} \frac{1}{q/t} = \frac{1}{q} \end{aligned}$$

independent of  $t$ , since  $-\partial_q N_1$  decreases. (By the mean value theorem,  $1 - N_1(q/t) = -\partial_q N_1(c)(q/t)$  for some  $c < q/t$ .) Now

$$\int_x^\infty \frac{e^{-qs}}{s} \mu_t(ds) \leq \frac{1}{x} \int_0^\infty e^{-qs} \mu_t(ds) \leq \frac{1}{x} \cdot \frac{1}{q} < \varepsilon$$

for  $x > 1/q\varepsilon$ ; this estimate controls the tail. It follows that

$$\int_1^\infty \frac{e^{-qs}}{s} \mu_t(ds) \rightarrow \int_1^\infty \frac{e^{-qs}}{s} \frac{ds}{2}$$

as  $t \rightarrow \infty$ . This finishes the proof.  $\square$

### 3. Models of domain coarsening in two and three dimensions

*Universal scaling behavior.* One of the classic dynamic scaling phenomena observed in material systems is Ostwald ripening, a process that occurs during the condensation of a supersaturated vapor, for example (think of clouds or fog), or during phase separation in metallic alloys. Many nuclei of the new phase appear and grow until there is a rough equilibrium with a complex arrangement of particles. The system continues to evolve in the late stages of this phase transition, however, driven by fluxes generated by curvature variation. In certain mixtures of metals, Ostwald observed that the typical particle size grows like  $t^{1/3}$ . The total phase fraction is conserved; large particles grow while small particles shrink and disappear.

An important paradigm for understanding this power-law scaling behavior is the model of Lifshitz and Slyozov [31] and Wagner [50] for the evolution of the particle size distribution. In this section we aim to describe how the LSW model fits in a hierarchy of multidimensional domain coarsening models similar to the one-dimensional hierarchy of the previous section, and indicate how recent mathematical analysis has helped to clarify why the LSW model is an unsatisfactory explanation for power-law scaling behavior as observed in practice.

*Diffuse and sharp interfaces.* We start with diffuse interface models (though there is research relating these to even more microscopic stochastic Ising models). Domain walls are “diffuse interfaces” which become “sharp interfaces” in the limit that their characteristic width divided by a macroscopic scale is taken to zero. The free energy concentrates on domain walls and becomes proportional to the interface surface area. Formally, curvature is the gradient of surface area (as we will see), and gradient flow means that coarsening in multidimensional systems is driven in many cases by the curvature of interfaces.

Allen & Cahn [3] argued by physical considerations that weakly curved domain walls in the multidimensional PDE

$$\partial_t u = -f(u) + \Delta u \tag{3.1}$$

should move with normal velocity  $v$  proportional to mean curvature  $\kappa$ . This result was derived a few years later within the systematic formal approximation procedure of matched asymptotic expansions by Rubinstein, Sternberg & Keller [48]. The same method was used to produce many interesting singular limits in the phase field system (a model of solidification) by Caginalp [8].

The hierarchy that leads to the LSW model of Ostwald ripening starts with the Cahn-Hilliard equation (a generalized diffusion equation)

$$\partial_t u = \Delta(f(u) - \varepsilon^2 \Delta u). \quad (3.2)$$

The sharp-interface limit of the Cahn-Hilliard equation turns out to be a model of phase transition kinetics found by Mullins & Sekerka to generate shape instabilities [46]: Given an interface set  $\Gamma$  separating two phases of material, steady-state diffusion of material (subject to a Gibbs-Thomson boundary condition for chemical potential) produces a jump in flux that drives the interface to move. Thus, the normal velocity  $v$  of  $\Gamma(t)$  is determined by solving a boundary value problem of the following form (setting physical constants to 1):

$$\Delta u = 0 \quad \text{in } \mathbb{R}^3 \setminus \Gamma, \quad (3.3)$$

$$u = \kappa \quad \text{on } \Gamma, \quad (3.4)$$

$$v = [\partial_\nu u]_-^+ \quad \text{on } \Gamma. \quad (3.5)$$

The LSW model arises from this model in a dilute regime in which particles are widely separated and the potential field  $u$  is approximated by a sum of “monopoles”  $\sum a_j/|x - x_j|$  plus a constant mean field. Below we will describe how the monopole model arises naturally by restriction from the gradient structure of the Mullins-Sekerka model. The LSW model then inherits a gradient structure that turns out to be useful for some things—see section 4.2 for an example.

### 3.1. Diffuse and sharp-interface models of nanoscale island coarsening

Rather than discuss the simplest situation in detail, it is interesting to consider a recent treatment of a problem of current interest related to nanoscale material interfaces. Understanding the dynamics of nanoscale structures is important for much emerging technology (such as the production of gem diamonds(!) by chemical vapor deposition [16]).

We study the motion of “steps” on surfaces of pure crystals in the vicinity of atomically flat (so-called *vicinal* surfaces). On the atomic level, material surfaces can be modeled discretely in terms of atoms occupying a lattice, or they can be modeled on the large scale at the continuum level by smooth surfaces when the surface is atomically rough. In an intermediate range described nicely in a review article of Jeong and Williams [26], crystalline materials can have atomically flat surfaces with partially filled layers of atoms on top. The edges of these layers (“steps”) are atomically rough and can be described by smooth curves that bound a raised “terrace” on the surface. Motion of these steps occurs due to thermal agitation of atoms along step edges and attachment and detachment of single atoms (adatoms) that diffuse on the terraces.

A classic continuum model of this step motion, that considers steps as smooth curves forming sharp interfaces between terraces, is the Burton-Cabrera-Franck (BCF) model [7]; see the treatment by Bales & Zangwill [5]. We will first describe the BCF model and then discuss a diffuse-interface approximation developed by Otto et al. [45] that yields the BCF model in the sharp-interface limit.

**BCF model.** The BCF model is based on the step-terrace description of the surface together with a number of assumptions: On the terraces, adatoms: (i) are deposited at rate  $F$  per site; (ii) desorb from the terrace with lifetime  $\tau$ ; and (iii) diffuse, hopping to neighboring sites with rate  $D$ . Let  $\rho(x, t)$  denote the adatom density (expected number of atoms at site  $x$ ) on a terrace (a region in  $\mathbb{R}^2$  bounded by a union of smooth curves). The processes can be accounted for by a discrete model. In a short time interval  $\Delta t$ , the adatom density at site  $x$  on the terrace changes according to

$$\rho(x, t + \Delta t) \sim \rho(x, t) + F\Delta t - \rho(x, t)\Delta t/\tau + D\Delta t \sum_{x' \in N(x)} (\rho(x', t) - \rho(x, t)),$$

where the sum goes over a set  $N(x)$  of neighbors of site  $x$ . If the lattice spacing is  $a$ , passing to a continuum model yields the PDE

$$\partial_t \rho(x, t) = Da^2 \Delta \rho + F - \rho/\tau \quad \text{in } \mathbb{R}^2 \setminus \Gamma \quad (3.6)$$

where  $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2$  is the Laplacian, and  $\Gamma$  is a set of curves comprising the steps.

Near steps, BCF suppose that adatoms attach to or detach from terrace edges at different rates from the upper and lower terraces. (This models an effect known as the Ehrlich-Schwoebel barrier effect— adatoms on the upper terrace experience a higher-than-usual potential barrier to get over

the edge of the step.) Attachment produces a net normal velocity  $v$  for the step in the direction of the lower terrace (denoted as the plus side here), and is driven by the difference between the terrace step density  $\rho(x, t)$  and an equilibrium constant  $\rho_*$  corrected by a term proportional to step curvature  $\kappa(x, t)$ :

$$\frac{v}{a} = k^+(\rho^+ - \rho_*(1 + \xi\kappa)) + k^-(\rho^- - \rho_*(1 + \xi\kappa)) \quad \text{on } \Gamma. \tag{3.7}$$

Here  $\rho^+(x, t)$  and  $\rho^-(x, t)$  are limits of  $\rho(y, t)$  as  $y$  approaches position  $x$  on the step from the lower and upper terraces respectively. The quantities  $k^+$ ,  $k_-$ ,  $\rho_*$  and  $\xi$  are constants.

The continuum model is completed by an equation that states that the flux of adatoms diffusing to the step edges balances the rate at which they attach to the step: Letting  $\partial_\nu$  denote the derivative along the unit normal  $\nu$  that points from the upper to the lower terrace,

$$Da \partial_\nu \rho^+ = k^+(\rho^+ - \rho_*(1 + \xi\kappa)), \tag{3.8}$$

$$-Da \partial_\nu \rho^- = k^-(\rho^- - \rho_*(1 + \xi\kappa)). \tag{3.9}$$

*Nondimensionalization and quasistatic limit.* We nondimensionalize time and space by  $t = T\tilde{t}$ ,  $x = L\tilde{x}$ . We take the time scale of interest to be  $T = 1/F$ , the mean rate that layers are deposited. We presume the lifetime  $\tau \gg T$  is long by comparison.

The length scale of interest  $L$  is taken to balance the effects that curvature and deposition have upon density variations—we require

$$\frac{\rho_*\xi}{L} \sim \frac{FL^2}{Da^2}, \quad \text{so} \quad L = \left( \frac{Da^2\rho_*\xi}{F} \right)^{1/3},$$

and scale excess density according to

$$w = \rho - \rho_* = \frac{\rho_*\xi}{L}\tilde{w}.$$

Then in terms of new variables (and dropping the tildes), the BCF model equations take the form

$$\Delta w + 1 = 0 \tag{3.10}$$

on terraces, and

$$v = \partial_\nu w^+ - \partial_\nu w^-, \tag{3.11}$$

$$\zeta_+ \partial_\nu w^+ = w^+ - \kappa, \tag{3.12}$$

$$-\zeta_- \partial_\nu w^- = w^- - \kappa \tag{3.13}$$

at steps.

The time derivative  $\partial_t w$  in (3.10) is neglected by supposing the diffusion time  $L^2/Da^2$  is small compared to  $T$ . When  $\zeta_{\pm} = 0$ , the conditions in (3.12)–(3.13) reduce to a Gibbs-Thomson boundary condition at steps. In this case, in the absence of deposition flux (replacing (3.10) by  $\Delta w = 0$ ), the interface dynamics reduces to the Mullins-Sekerka model.

**Diffuse-interface approximation.** We now describe the diffuse-interface approximation to this BCF model that was constructed by Otto, Penzler, Rätz, Rump, and Voigt [45]. The height of terraces in the sharp-interface model is an integer times the thickness of an atomic layer (taken as 1 here). In the diffuse approximation, the material surface height  $z(x, t)$  is modeled as a smooth function whose “free energy”

$$\mathcal{F}(z) = \int_{\mathbb{R}^2} \frac{\varepsilon}{2} |\nabla z|^2 + \frac{1}{\varepsilon} G(z) \, dx$$

is bounded. Here  $G : \mathbb{R} \rightarrow \mathbb{R}$  is a smooth function that we take to be like  $\sin^2(\pi z)$ : periodic with period 1, zero at integers, and positive otherwise. We take it normalized so  $\int_0^1 \sqrt{2G(z)} \, dz = 1$ . If  $\mathcal{F}(z)$  is small, it should mean  $z$  is close to an integer in large regions corresponding to terraces, separated by narrow transition zones between.

The equation governing the evolution of  $z(x, t)$  will take the form of a modified Cahn-Hilliard equation. This takes the form of a diffusion equation

$$\partial_t z + \nabla \cdot \mathbf{j} = 1 \tag{3.14}$$

to hold everywhere in space, with the “flux”  $\mathbf{j}$  given in terms of a “chemical potential”  $w$  according to

$$(1 + \varepsilon^{-1} \zeta_2(z)) \mathbf{j} = -\nabla w, \tag{3.15}$$

$$w = \varepsilon \zeta_1(z) \partial_t z - \varepsilon \Delta z + \varepsilon^{-1} G'(z). \tag{3.16}$$

The drag coefficient  $\zeta_1(z)$  and mobility coefficient  $\zeta_2(z)$  are non-negative and 1-periodic. We take  $\zeta_2(z)$  to vanish on integers and be positive otherwise (like  $G$ ); this is to force the flux  $\mathbf{j}$  to be small in the transition layers as  $\varepsilon \rightarrow 0$ .

Equations (3.14)–(3.16) comprise the OPRRV model. For a closed system (replace 1 by 0 in (3.14)) with no-flux boundary conditions  $\nu \cdot \mathbf{j} = 0$ , the “free energy”  $\mathcal{F}(z)$  decreases:

$$\begin{aligned} \frac{d}{dt} \mathcal{F}(z) &= \frac{d}{dt} \int \left( \frac{\varepsilon}{2} |\nabla z|^2 + \frac{1}{\varepsilon} G(z) \right) \, dx = \int (-\varepsilon \Delta z + \varepsilon^{-1} G'(z)) \, \partial_t z \\ &= \int (w - \varepsilon \zeta_1 \partial_t z) \partial_t z = \int \nabla w \cdot \mathbf{j} - \varepsilon \zeta_1 (\partial_t z)^2 \leq 0. \end{aligned}$$

*Method of matched asymptotic expansions.* We describe how the evolution of  $z$  by motion of transition layers yields the BCF interface dynamics formally in the limit  $\varepsilon \rightarrow 0$ , using the method of matched asymptotic expansions. Mathematically, the method involves constructing an *approximate solution* to the OPRRV equations in two overlapping zones: One constructs (i) an inner expansion to be valid distances  $\mathcal{O}(\varepsilon)$  from the evolving steps  $\Gamma(t)$ ; and (ii) an outer expansion to be valid distances  $\mathcal{O}(1)$  from the steps. The two expansions are linked by a matching procedure at intermediate distances from the steps. Maximizing the order of accuracy of the approximation will require step evolution to be given by the sharp-interface BCF model. Further insight on the formal matching procedure can be found in [20, 8].

We are mainly interested in interface dynamics, so we will neglect the deposition flux, replacing 1 by 0 on the right hand sides of (3.10) and (3.14).

Consider the outer expansion first: We seek  $z(x, t)$  in the form

$$z(x, t) = z_0(x, t) + \varepsilon z_1(x, t) + \mathcal{O}(\varepsilon^2),$$

where  $z_0, z_1$  are independent of  $\varepsilon$ , with similar expansions for  $w$  and  $\mathbf{j}$ . Plugging these into the equations, we require the  $\mathcal{O}(\varepsilon^{-1})$  terms balance, yielding

$$G'(z_0) = 0 \quad \text{and} \quad \zeta_2(z_0)\mathbf{j}_0 = 0.$$

By our hypotheses, the first equation forces  $z_0(x, t)$  to take integer values, constant in components complementary to the steps  $\Gamma(t)$ . Then  $\zeta_2(z_0) = 0$  and  $\mathbf{j}_0$  is not restricted.

Balancing terms of order  $\mathcal{O}(1)$  yields three equations:

$$\begin{aligned} \partial_t z_0 + \nabla \cdot \mathbf{j}_0 &= 0, \\ (1 + \zeta_2'(z_0)z_1)\mathbf{j}_0 + \zeta_2(z_0)\mathbf{j}_1 &= -\nabla w_0, \\ w_0 &= G''(z_0)z_1. \end{aligned}$$

Since  $0 = \partial_t z_0 = \zeta_2(z_0) = \zeta_2'(z_0)$ , this simplifies to  $\mathbf{j}_0 = -\nabla w_0$  and

$$-\Delta w_0 = 0, \tag{3.17}$$

with  $w_0 = G''(0)z_1$ . This is all we need from the outer expansion.

Next consider the inner expansion, distances  $\mathcal{O}(\varepsilon)$  from the steps  $\Gamma(t)$ , which we take to be a union of smooth curves independent of  $\varepsilon$ . We suppose without loss that the step models a single atomic layer, with  $z_0^+ = 0, z_0^- = 1$ . We will need to “stretch” the coordinates normal to the steps. In

the neighborhood of some point on  $\Gamma$ , introduce the signed distance

$$\Phi(x, t) := \pm \text{dist}(x, \Gamma(t))$$

from  $x$  to  $\Gamma(t)$ , where we take the sign as  $+$  on the lower terrace and  $-$  on the upper. Then  $\nu(x, t) = \nabla\Phi(x, t)$  is a unit vector normal to  $\Gamma$  if  $x \in \Gamma$ . The normal velocity of  $\Gamma$  at  $x$  is

$$v(x, t) = -\partial_t\Phi.$$

We let  $r = \Phi/\varepsilon$ , and for  $x$  near  $\Gamma$  change variables via  $x = y + \varepsilon r\nu(y, t)$  for  $y \in \Gamma(t)$ . In principle, this yields a local change of variables  $(x, t) \mapsto (y, r, t)$ . It is a convenient rule, however, to regard quantities  $Q(y, r, t)$  as having the form  $Q(x, r, t)$ , in which  $x$  is not restricted to lie on  $\Gamma$  but  $Q$  must be constant as  $x$  varies along  $\nu$  in the first argument, i.e.,  $\nu \cdot \nabla_x Q = 0$ .

We seek our inner expansion in the form

$$\begin{aligned} z &= Z_0(x, r, t) + \varepsilon Z_1(x, r, t) + \mathcal{O}(\varepsilon^2), \\ \mathbf{j} &= J_0(x, r, t) \nu(x, t) + \mathcal{O}(\varepsilon), \\ w &= W_0(x, r, t) + \mathcal{O}(\varepsilon), \quad r = \Phi(x)/\varepsilon. \end{aligned}$$

Here, the indicated functions of  $(x, r, t)$  are to be independent of  $\varepsilon$  and need to be determined for all  $r \in (-\infty, \infty)$ . Evaluating derivatives by the chain rule, we have

$$\begin{aligned} \partial_t z &= -\varepsilon^{-1} v \partial_r Z_0 + \mathcal{O}(1), \\ \nabla \cdot \mathbf{j} &= \varepsilon^{-1} \nabla\Phi \cdot \nu \partial_r J_0 + \mathcal{O}(1) = \varepsilon^{-1} \partial_r J_0 + \mathcal{O}(1), \\ \nabla w &= \varepsilon^{-1} \nabla\Phi \partial_r W_0 + \mathcal{O}(1), \\ \Delta z &= \varepsilon^{-2} \partial_r^2 Z_0 + \varepsilon^{-1} \Delta\Phi \partial_r Z_0 + \mathcal{O}(1). \end{aligned}$$

A fact that must be left to the reader to check is that for  $x \in \Gamma(t)$ ,

$$\Delta\Phi(x, t) = \kappa(x, t)$$

is the curvature of  $\Gamma(t)$ .

We now use these expressions in the OPRRV model equations (3.14)–(3.16) and match terms of order  $\mathcal{O}(\varepsilon^{-1})$  to find:

$$-v \partial_r Z_0 + \partial_r J_0 = 0, \tag{3.18}$$

$$-\partial_r^2 Z_0 + G'(Z_0) = 0, \tag{3.19}$$

$$\zeta_2(Z_0) J_0 = -\partial_r W_0. \tag{3.20}$$

From these equations it follows that

$$-v Z_0 + J_0 = \lambda(x, t) \tag{3.21}$$

independent of  $r \in (-\infty, \infty)$ . We match to the outer expansion in the regime where  $r = \Phi/\varepsilon$  is large but  $\Phi$  is small. This leads us to require

$$Z_0(r) \rightarrow 1 \text{ as } r \rightarrow -\infty, \quad 0 \text{ as } r \rightarrow \infty, \tag{3.22}$$

and we satisfy (3.19) by taking  $Z_0 = Z_0(r)$  to be a *domain wall* independent of  $x, t$  and centered so that  $Z_0(0) = 1/2$ , say. From (3.19) we find  $Z_0$  satisfies

$$\partial_r Z_0 = -\sqrt{2G(Z_0)}, \quad Z_0(0) = 1/2. \tag{3.23}$$

Taking  $r \rightarrow +\infty$  then  $-\infty$  in (3.21) now leads to the matching conditions

$$\begin{aligned} \lambda &= J_0(x, +\infty, t) = \nu \cdot \mathbf{j}_0^+ = -\partial_\nu w_0^+, \\ v &= J_0(x, -\infty, t) - \lambda = \partial_\nu w_0^+ - \partial_\nu w_0^-. \end{aligned}$$

At this point, equations (3.17) and (3.24) correspond respectively to (3.10) and (3.11), and it remains to recover the equations (3.12)–(3.13) that govern the attachment kinetics. From terms of order  $\mathcal{O}(1)$  we need only observe that

$$W_0 = \zeta_1(Z_0)(-v \partial_r Z_0) - \kappa \partial_r Z_0 - \partial_r^2 Z_1 + G''(Z_0)Z_1. \tag{3.24}$$

Since  $(-\partial_r^2 + G''(Z_0))\partial_r Z_0 = 0$ , we can say

$$\int_{-\infty}^{\infty} (\partial_r Z_0)(-\partial_r^2 + G''(Z_0))Z_1 \, dr = 0.$$

Thus, a *necessary condition* for the solvability of (3.24) for  $Z_1$  is that

$$\int_{-\infty}^{\infty} (\partial_r Z_0)(W_0 + \kappa \partial_r Z_0 + v\zeta_1 \partial_r Z_0) \, dr = 0.$$

Now, using (3.20) and (3.21) we compute

$$\begin{aligned} \kappa \int_{-\infty}^{\infty} (\partial_r Z_0)^2 + v \int_{-\infty}^{\infty} \zeta_1 (\partial_r Z_0)^2 &= - \int_{-\infty}^{\infty} W_0 (\partial_r Z_0) \, dr \\ &= -[W_0 Z_0]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (\zeta_2 J_0) Z_0 = w_0^- - v \int_{-\infty}^{\infty} \zeta_2 Z_0^2 - \lambda \int_{-\infty}^{\infty} \zeta_2 Z_0, \end{aligned}$$

and

$$w_0^+ - w_0^- = \int_{-\infty}^{\infty} \partial_r W_0 \, dr = - \int_{-\infty}^{\infty} \zeta_2 J_0 = -v \int_{-\infty}^{\infty} \zeta_2 Z_0 - \lambda \int_{-\infty}^{\infty} \zeta_2.$$

By (3.24) and (3.24), and since  $G$  is normalized so that

$$\int_0^1 \sqrt{2G(z)} \, dz = \int_{-\infty}^{\infty} (\partial_r Z_0)^2 \, dr = 1,$$

this yields the BCF equations

$$\zeta_+ \partial_\nu w_0^+ = w_0^+ - \kappa, \tag{3.25}$$

$$-\zeta_- \partial_\nu w_0^- = w_0^- - \kappa, \tag{3.26}$$

provided the following hold:

$$\zeta_- = \int_{-\infty}^{\infty} (\zeta_1(Z_0)(\partial_r Z_0)^2 + \zeta_2(Z_0)Z_0^2) dr = \int_{-\infty}^{\infty} \zeta_2(Z_0)Z_0 dr, \tag{3.27}$$

$$\zeta_+ = \int_{-\infty}^{\infty} \zeta_2(Z_0)(1 - Z_0) dr. \tag{3.28}$$

By changing variables using (3.23), these constraints can be written in the form

$$\zeta_- = \int_0^1 \left( \zeta_1(z)\sqrt{2G(z)} + \frac{\zeta_2(z)z^2}{\sqrt{2G(z)}} \right) dz = \int_0^1 \frac{\zeta_2(z)z}{\sqrt{2G(z)}} dz, \tag{3.29}$$

$$\zeta_+ = \int_0^1 \frac{\zeta_2(z)(1 - z)}{\sqrt{2G(z)}} dz. \tag{3.30}$$

This completes the formal derivation. We point out that deriving the Mullins-Sekerka sharp-interface model (3.3)–(3.5) from the Cahn-Hilliard equation (3.2) is just a special case of the above, taking  $\zeta_1 = \zeta_2 = 0$ . A rigorous justification of the sharp-interface limit in this case was given by Alikakos, Bates, and Chen [1]. For work relating the Cahn-Hilliard equation to the LSW model, see the recent paper [2] and the references therein.

### 3.2. Gradient structure for Mullins-Sekerka flow

For use below, here we explain (essentially following Niethammer & Otto [39]) how the Mullins-Sekerka model can be described formally as gradient flow for surface area with respect to a certain metric structure on a “manifold” of smooth surfaces in  $\mathbb{R}^3$ . (We remark that for the closely related problem of Hele-Shaw flow between parallel plates, this type of gradient flow structure, as described by Otto [44], was exploited by Glasner [24] to derive a corresponding diffuse-interface model in a very interesting manner.)

First, recall the general structure of gradient flow for an energy functional  $\mathcal{E} : \mathcal{M} \rightarrow \mathbb{R}$  where  $\mathcal{M}$  is a Riemannian manifold with metric  $g$ : A solution trajectory is a curve  $t \mapsto z(t) \in \mathcal{M}$  with tangent vector  $\partial_t z \in T_{z(t)}\mathcal{M}$  that satisfies

$$g_{z(t)}(\partial_t z, v) = -d\mathcal{E}(z(t))v \quad \text{for all } v \in T_{z(t)}\mathcal{M}.$$

Now, we consider  $\mathcal{M}$  to be the “manifold” of smooth bounded surfaces  $\Gamma$  that are boundaries of bounded domains in  $\mathbb{R}^3$ . Formally, elements of the tangent space  $T_\Gamma\mathcal{M}$  correspond to “normal velocity fields”  $v : \Gamma \rightarrow \mathbb{R}$ . We define a “metric” on  $T_\Gamma\mathcal{M}$  as follows: To each such  $v \in T_\Gamma\mathcal{M}$  we associate a harmonic “potential”  $u = \mathcal{T}v : \mathbb{R}^3 \rightarrow \mathbb{R}$  by solving<sup>b</sup> the PDE boundary value problem

$$\begin{aligned} \Delta u &= 0 \text{ in } \mathbb{R}^3 \setminus \Gamma, \\ \nu \cdot \nabla u^+ - \nu \cdot \nabla u^- &= -v \text{ on } \Gamma, \\ u(x) &\rightarrow 0 \text{ as } |x| \rightarrow \infty. \end{aligned}$$

(Here  $\nu$  is the unit outward normal to the domain enclosed by  $\Gamma$ , and  $u^+, u^-$  are respective limits on  $\Gamma$  along  $\nu$  from the outside and inside respectively.) Given  $v_1, v_2 \in T_\Gamma\mathcal{M}$ , let  $u_1 = \mathcal{T}v_1, u_2 = \mathcal{T}v_2$  and put

$$g_\Gamma(v_1, v_2) = \int_{\mathbb{R}^3} \nabla u_1 \nabla u_2 = \int_\Gamma u_1 [n \nabla u_2]_+^- = \int_\Gamma u_1 v_2 = \int_\Gamma v_1 v_2 \quad (3.31)$$

We let  $\mathcal{E}(\Gamma)$  be the surface area of  $\Gamma$ . It is known that if  $\Gamma(t)$  is smoothly evolving with normal velocity  $v(t)$ , then

$$\frac{d}{dt} \mathcal{E}(\Gamma(t)) = \int_\Gamma \kappa v =: d\mathcal{E}(\Gamma)v,$$

where  $\kappa$  is the sum of principal curvatures of  $\Gamma$  (positive for spheres), and the volume of the domain  $\Omega(t)$  enclosed by  $\Gamma$  evolves by

$$\frac{d}{dt} \text{vol}(\Omega(t)) = \int_\Gamma v.$$

Mullins-Sekerka flow is gradient flow for surface area with *enclosed volume conserved*. Let  $\mathcal{M}_0$  be a submanifold of  $\mathcal{M}$  corresponding to surfaces with *fixed* enclosed volume. Velocity fields  $v \in T_\Gamma\mathcal{M}_0$  should satisfy  $\int_\Gamma v = 0$ . Gradient flow requires that  $\Gamma(t)$  evolves so that its normal velocity  $v$  satisfies

$$g_\Gamma(v, \tilde{v}) = -d\mathcal{E}(\Gamma)\tilde{v}, \quad \text{or} \quad \int_\Gamma u\tilde{v} = - \int_\Gamma \kappa\tilde{v}, \quad \text{for all } \tilde{v} \in T_\Gamma\mathcal{M}_0.$$

---

<sup>b</sup>How to do this technically: Let  $\mathcal{H} = \{u \in L^6(\mathbb{R}^3) \mid \nabla u \in L^2(\mathbb{R}^3)^3\}$  be a Hilbert space with  $\|u\|_{\mathcal{H}} = (\int |\nabla u|^2)^{1/2}$ .  $\mathcal{H}$  is complete due to a critical Sobolev embedding theorem. Find  $u$  so

$$\langle u, v \rangle_{\mathcal{H}} = \int_{\mathbb{R}^3} \nabla u \cdot \nabla w = \int_\Gamma v w \quad \text{for all } w \in \mathcal{H}.$$

Since this holds for all smooth  $\tilde{v}$  with  $\int_{\Gamma} \tilde{v} = 0$ , we can infer that  $u = -\kappa + \theta(t)$  on  $\Gamma(t)$ , where  $\theta(t)$  is constant in space. This yields the Mullins-Sekerka motion law, since we have

$$\begin{aligned} \Delta(-u + \theta) &= 0 \quad \text{in } \mathbb{R} \setminus \Gamma(t), \\ -u + \theta &= \kappa \quad \text{on } \Gamma, \\ v &= [\nu \cdot \nabla(-u + \theta)]_{-}^{+} \quad \text{on } \Gamma. \end{aligned}$$

Note that  $\theta = \lim_{|x| \rightarrow \infty} (-u + \theta)$  is the “mean field,” the limit at infinity of the harmonic function  $-u + \theta$  determined by the curvature of  $\Gamma(t)$ .

### 3.3. Monopole models by restricted gradient flow of surface energy

The morphology of domains coarsening according to Mullins-Sekerka flow can be complex. Singularities may occur frequently through the shrinking of small blobs to zero size, or the “pinch-off” of necks in dumb-bell shaped regions, for example. When the minority phase occupies a small fraction of a sample region, however, frequently the morphology seen is that of a dilute suspension of approximately spherical domains. (Presumably one sees spheres due to some sort of local minimization of surface area with constrained enclosed volume).

This leads us to consider a geometrically simplified model in which the evolving surface  $\Gamma$  is constrained to consist of a collection of spheres. As we show below following S. Dai’s Ph.D. thesis [14], Mullins-Sekerka gradient flow constrained geometrically to spheres exactly yields the classic *monopole model*, in which the harmonic potential (diffusion field)  $u = \mathcal{T}v$  is a superposition of monopole fields  $1/|x - x_i|$ . The monopole model is important due to the fact that it is amenable to large-scale simulation; computations involving hundreds of thousands of spheres have been performed (taking some shortcuts).

We restrict attention to a submanifold  $\mathcal{M}_N$  of surfaces  $\Gamma$  consisting of a collection of spheres  $\Gamma_i$  bounding a fixed number  $N > 0$  of non-overlapping balls  $B_i$  where  $|x - x_i| < R_i$ ,  $i = 1, \dots, N$ . We consider the centers  $x_i \in \mathbb{R}^3$  fixed, and fix total volume, so  $\sum R_i^3 = Q$  is constant. The manifold  $\mathcal{M}_N$  is  $N - 1$  dimensional and each tangent vector  $v \in T_{\Gamma} \mathcal{M}_N$  corresponds to a normal velocity field with  $v = v_i = \dot{R}_i$  constant on the sphere  $\Gamma_i$  where  $|x - x_i| = R_i$ .

We will find  $u = \mathcal{T}v$  explicitly as a superposition of monopoles:

$$u(x) = \begin{cases} \sum_{j=1}^N \frac{a_j}{|x - x_j|}, & \text{if } |x - x_j| \geq R_j \text{ for all } j, \\ \sum_{j \neq i} \frac{a_j}{|x - x_j|} + \frac{a_i}{R_i}, & \text{if } |x - x_i| < R_i. \end{cases}$$

Evidently  $\Delta u = 0$  in  $\mathbb{R}^3 \setminus \Gamma$  and  $u(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ . The jump condition on  $\Gamma_i$  reads  $[\nu \cdot \nabla u]_{-}^{+} = -a_i/R_i^2 = -v_i$ , hence we must have

$$a_i = R_i^2 v_i,$$

and consequently,

$$\sum a_i = 0.$$

To describe how the  $v_i = \dot{R}_i$  are determined, we must show how the monopole amplitudes  $a_i$  can be determined from the sphere radii  $R_i$  by solving a linear system of equations. Given a tangent vector  $\tilde{v} \in T_{\Gamma} \mathcal{M}_N$ , we compute

$$\begin{aligned} g_{\Gamma}(v, \tilde{v}) &= \int_{\Gamma} u \tilde{v} = \sum_i \int_{\Gamma_i} \tilde{v}_i \sum_j \frac{a_j}{|x - x_j|} \\ &= \sum_i 4\pi R_i^2 \tilde{v}_i \left( \frac{a_i}{R_i} + \sum_{j \neq i} \frac{a_i}{|x_i - x_j|} \right), \end{aligned}$$

since by the mean value property of harmonic functions ( $\Delta(1/|x|) = 0$ ),

$$\int_{\Gamma_i} \frac{1}{|x - x_j|} = \frac{4\pi R_i^2}{|x_i - x_j|}.$$

For the surface area  $\mathcal{E} = \sum 4\pi R_i^2$  we can write

$$d\mathcal{E}(\Gamma)\tilde{v} = \sum 8\pi R_i \tilde{v}_i = \sum 4\pi R_i^2 \left( \frac{2}{R_i} \tilde{v}_i \right) = \int_{\Gamma} \kappa \tilde{v}.$$

Hence, gradient flow ( $g_{\Gamma}(v, \tilde{v}) = -d\mathcal{E}(\Gamma)\tilde{v}$  for all  $\tilde{v} \in T_{\Gamma} \mathcal{M}_N$ ) means that

$$\sum_i 4\pi R_i^2 \tilde{v}_i \left( \frac{2}{R_i} + \frac{a_i}{R_i} + \sum_{j \neq i} \frac{a_i}{|x_i - x_j|} \right) = 0 \quad \text{whenever} \quad \sum_i R_i^2 \tilde{v}_i = 0.$$

Hence the term in parentheses must be independent of  $i$ , and we denote it by  $\tilde{\theta}(t)$ . We now have  $N + 1$  equations to determine  $N + 1$  unknowns

$v_1, \dots, v_N$  and  $\tilde{\theta}$ , namely:

$$R_i^3 v_i + \sum_{j \neq i} \frac{R_i^2 R_j^2 v_j}{|x_i - x_j|} = R_i^2 \left( \tilde{\theta} - \frac{2}{R_i} \right), \quad \sum_i R_i^2 v_i = 0. \quad (3.32)$$

In matrix-vector form this reads:

$$A\vec{v} = \tilde{\theta}\vec{s} - \vec{b}, \quad \vec{s} \cdot \vec{v} = 0,$$

with  $s_i = R_i^2$ ,  $b_i = R_i^2(\tilde{\theta} - 2/R_i)$ . The matrix  $A$  is symmetric and positive definite since

$$\vec{v} \cdot A\vec{v} = \frac{g\Gamma(v, v)}{4\pi} > 0.$$

The solution can be expressed in the form

$$\vec{v} = \tilde{\theta}A^{-1}\vec{s} - A^{-1}\vec{b}, \quad \tilde{\theta} = \frac{\vec{s} \cdot A^{-1}\vec{b}}{\vec{s} \cdot A^{-1}\vec{s}}.$$

This completes the derivation of the monopole model:

$$\partial_t R_i = v_i(\vec{R}_i).$$

This equation applies up to a time when one or more particle radii vanish ( $R_i \rightarrow 0$ ), after which the system is continued with fewer particles, or until particles collide.

An advantage of this derivation of the monopole model by restricted gradient flow is that it shows that *the velocities  $\dot{R}_i$  in the monopole model are well-defined as long as the balls  $B_i$  are non-overlapping*. Moreover, the derivation shows how the velocities can be determined in terms of a positive definite matrix; this apparently has not been recognized before and could be useful in numerical computations.

### 3.4. Lifshitz-Slyozov-Wagner mean-field model

The first quantitative explanation of the  $t^{1/3}$  power-law growth of typical domain size observed during Ostwald ripening was provided by work of Lifshitz and Slyozov [31] and Wagner [50], based upon arguments involving self-similar behavior for a Liouville equation governing the particle size distribution in a regime where all the interactions of particles are subsumed in one mean-field coupling term. In this subsection we describe this LSW model and indicate how recent rigorous analysis has clarified the fact that the self-similar nature of experimentally observed size distributions is not completely explained by the mean-field model, and must depend upon factors not taken into account by it.

In order to derive the LSW model, we look at at the dilute limit in which the typical particle radius  $R_i$  is very small compared to the interparticle distances  $|x_i - x_j|$ . Replacing  $x_i$  by  $\bar{x}x_i$  and letting  $\bar{x} \rightarrow \infty$ , the off-diagonal terms in the system vanish ( $A \rightarrow \text{diag}(R_i^3)$ ), and we are left with the system

$$R_i^3 v_i = R_i^2 \left( \tilde{\theta} - \frac{2}{R_i} \right),$$

that is,

$$\partial_t R_i = \frac{1}{R_i} \left( \tilde{\theta} - \frac{2}{R_i} \right).$$

Since  $\sum R_i^2 \partial_t R_i = 0$ , the mean field  $\tilde{\theta}$  is determined by

$$\tilde{\theta} = \frac{\sum 2}{\sum R_i} = \frac{2}{R_{\text{av}}},$$

where  $R_{\text{av}} = \sum R_i / N$  is the average particle radius.

For later reference, we note that the metric  $g_\Gamma$  degenerates to diagonal form and the surface area decays as follows: Using that  $\sum 4\pi R_i^2 v_i = 0$ ,

$$\frac{dE}{dt} = \sum 8\pi R_i v_i = \sum 4\pi R_i^2 \left( \frac{2}{R_i} - \tilde{\theta} \right) v_i = -4\pi \sum R_i^3 v_i^2 = -g_\Gamma(v, v).$$

Coarsening proceeds according to the LSW model in the following way. Particles larger than average grow, and ones smaller than average shrink: If  $R_i > R_{\text{av}}$  then  $\partial_t R_i > 0$ ; if  $R_i < R_{\text{av}}$  then  $\partial_t R_i < 0$ . Very small particles will vanish in a finite time: If the smallest particle is the  $i$ th and  $R_i$  is small, then  $\partial_t R_i^3 \sim -6$ , and  $R_i^3 \sim 6(T_i - t_i)$  where  $R_i(T_i) = 0$ . Beyond this time the system continues with fewer particles.

**Theory of Lifshitz & Slyozov and Wagner.** The arguments of LSW to explain  $t^{1/3}$  growth of typical particle size involve writing a Liouville (or Fokker-Planck) equation for the size distribution function. We set

$$\varphi(R, t) = \frac{\# \text{ of particles of radius } \geq R}{\# \text{ at } t = 0}.$$

Since particle radii satisfy

$$\partial_t R = V(R, t) := \frac{1}{R} \left( \tilde{\theta} - \frac{2}{R} \right), \tag{3.33}$$

and this ODE preserves the order of particle sizes, conservation of particles implies  $\varphi$  satisfies a PDE:

$$\frac{d}{dt} \varphi(R(t), t) = \partial_t \varphi + V(R, t) \partial_R \varphi = 0. \tag{3.34}$$

In terms of a *number density*  $n(R, t)$  for the size distribution, we have

$$\varphi(R, t) = \int_R^\infty n(s, t) ds,$$

and the governing equation is written as a conservation law,

$$\partial_t n + \partial_R(V(R, t)n) = 0. \quad (3.35)$$

This is the LSW mean-field model. (For rigorous derivations of such PDE models from monopole models, see [37, 38, 39].) Conservation of total volume means that  $Q = \int_0^\infty R^3 n(R, t) dR$  is constant and leads to

$$\tilde{\theta}(t) = \frac{2 \int n dR}{\int R n dR} = \frac{2}{R_{\text{av}}(t)}.$$

*Scaling and self-similarity.* The particle growth law (3.33) has the scaling invariance  $R = \lambda R'$ ,  $t = \lambda^3 t'$ . If the solution of (3.35) achieves scale invariant form, so that

$$n(R, t) = \lambda^p n(\lambda R, \lambda^3 t),$$

then we must have  $p = 4$ , since

$$Q = \int R^3 n(R, t) dR = \int R^3 \lambda^p n(\lambda R, \lambda^3 t) dR = \lambda^{p-3-1} \int R'^3 n(R', t') dR'.$$

Thus, a scale-invariant solution should take the self-similar form

$$n(R, t) = t^{-4/3} F(R/t^{1/3}).$$

Now one has the questions: Can we expect to see this? And what should  $F$  be?

The equation actually admits a one-parameter family of self-similar solutions. Lifshitz & Slyozov argued as follows to explain that  $F$  should have a particular explicit form that is smooth with bounded support. Change variables, normalizing radius by its average, and introduce

$$\rho = \frac{R}{R_{\text{av}}}, \quad \tau = \log R_{\text{av}}^3.$$

Then  $\partial_t R^3 = 6(\rho - 1) = (\partial_t R_{\text{av}}^3)(\rho^3 + \partial_\tau \rho^3)$ , and we get

$$\partial_\tau \rho^3 = \gamma(\rho - 1) - \rho^3 \quad (3.36)$$

with  $\gamma = \gamma(\tau) = 6/(\partial_t R_{\text{av}}^3)$ . Expecting the system to settle into a self-similar regime, we expect  $\gamma(t) \rightarrow \gamma_\infty$ . This means that for large  $t$ ,

$$R_{\text{av}}^3 \sim ct, \quad \text{with } c = 6/\gamma_\infty.$$

One argues that  $\gamma_\infty = \frac{27}{4}$  (meaning  $c = 8/9$ , which differs from the classic  $4/9$  only due to the factor 2 in (3.33)) based on the dynamics of (3.36): If  $\gamma_\infty$  is smaller, then the right-hand side is everywhere negative and  $\rho$  becomes zero in finite time for all particles, contradicting total volume conservation. If  $\gamma_\infty$  is larger, then (3.36) admits two equilibria, a stable  $\rho_2$  above an unstable  $\rho_1$ . Self-similar solutions having bounded support exist in this case, but LSW argue they are unstable. Presuming some fraction of particle sizes lie above the unstable  $\rho_1$  means their  $\rho$  will approach  $\rho_2$  as  $t$  becomes large and this leads to growth of total volume, again contradicting conservation.

**Non-self-similar behavior.** These arguments of LSW are mathematically nonrigorous, but are physically precise and plausible and sparked a great deal of activity to investigate the predictions. It is fair to say the LSW analysis served as a paradigm for a range of related problems in materials science and solid state physics. Certain difficulties dogged the theory, however, especially the facts that observed size distributions are always broader than the predicted one, and that the assumption that the system is sufficiently dilute is never satisfied in real systems.

Rather recently, several groups of investigators [23, 12, 40] came to understand that solutions of mean-field LSW models such as (3.35) *need not* exhibit universal self-similar behavior, meaning perhaps that the LSW model lacks some feature of experimental systems which leads to observed self-similarity. A basic explanation is based upon equation (3.34). Evolution under this equation simply *stretches* the graph of  $\varphi$  according to the solution map  $R(0) \mapsto R(t)$  for the characteristic equation (3.33). This map is smooth (analytic, in fact), and thus for finite time it produces a smooth distortion of the initial distribution function.

Suppose the system initially has a maximal particle size  $R_*$ , so that initially  $\varphi(R, 0) = 0$  on  $(R_*, \infty)$  (maximal). The detailed way in which  $\varphi(R, 0)$  vanishes near the end of support will be qualitatively preserved for finite time. Niethammer & Pego [40] proved two facts for (3.34):

- (1) If there exist  $c > 0$  and  $p > 0$  such that initially

$$\varphi(R_* - r, 0) \geq cr^p,$$

then the rescaled solution *cannot* approach LSW's self-similar solution as  $t \rightarrow \infty$ .

- (2) A *necessary* condition for the solution to approach *some* self-similar solution as  $t \rightarrow \infty$  is that the initial distribution is “almost power-law”

near the maximal particle size. Namely,

$$\varphi(R_* - r, 0) \sim r^p L(r)$$

as  $r \rightarrow 0^+$  for some  $p \geq 0$  and some function  $L$  slowly varying near zero. (Recall this means  $\lim_{t \rightarrow 0} L(tx)/L(t) = 1$  for all  $x > 0$ .)

The condition in (2) says that  $r \mapsto \varphi(R_* - r, 0)$  is *regularly varying* at 0 with exponent  $p$ . This is the condition that figures in the necessary and sufficient conditions of the Tauberian theorem [19] and the rigidity lemma for scaling limits. Entire books have been written about it due to its importance in analysis and probability theory [6, 49]. In section 5 below we will show that regular variation is key for obtaining necessary and sufficient conditions to classify domains of attraction for self-similar behavior in Smoluchowski's coagulation equations.

**Comments on analysis.** Without getting into technical details, it is interesting to note that physical and structural considerations have some important consequences for the rigorous mathematical analysis of the LSW model. In the form (3.35) the model takes the form of a PDE conservation law as one has in shock wave theory. However, it is unwise to look to shock-wave theory for an appropriate topology to study the well-posedness of the initial-value problem. Instead, an appropriate topology that is physically meaningful should make it “difficult” to create large particles from nothing. On the other hand, small changes in particle size should be “easy,” even if the size distribution is highly peaked like a Dirac mass. Mathematically, one would ideally like to allow arbitrary size distributions, which means arbitrary probability measures after normalization by number.

Looking back at (3.34), it is rather more natural to regard particle size as the actively evolving dependent variable, and describe the size distribution by inverting the map  $R \mapsto \varphi$ , regarding  $R$  as a function of  $\varphi \in [0, 1]$ , the initial particle “rank.” The equation of evolution is just (3.33). The sup-norm distance between particle rankings corresponds to a transport metric called the  $L^\infty$  *Wasserstein distance* between probability distributions. (This is the minimal maximum size change needed to change one distribution to the other.) Based on this topology (in terms of volume  $v = R^3$ ), well-posedness of the LSW initial-value problem was proved in [41], with size distribution allowed to be an arbitrary probability distribution of compact support. Also, computations are best based on (3.33) rather than (3.35). For this model, it is much easier to attain high accuracy for long times following characteristics than with shock-capturing schemes. Over many years, a good number of investigators computing solutions numerically from (3.35)

determined (erroneously, as it turned out) that solutions always approached the smooth LSW form.

The restriction of compact support was removed in [42] using a gradient flow structure that the LSW model inherits from the (discrete) monopole model (despite the fact that (3.35) is a first-order PDE). The dissipation identity associated with this structure provides a compactness property useful for establishing the existence of solutions for any initial size distribution that is a probability measure giving finite expected volume.

#### 4. Rigorous power-law bounds on coarsening rates — The Kohn-Otto method

The scientific achievement of Lifshitz and Slyozov and Wagner in producing an explanation for the  $t^{1/3}$  power-law behavior for typical domain size in coarsening by phase separation was considerable; they spawned a large related literature that continues to expand rapidly. The LSW mean-field theory, however, is based upon a quite restrictive set of physical assumptions. In particular, the minority phase must be extremely dilute, and the particles spherical. These assumptions naturally fail in practical situations in metallic alloys with comparable phase fractions and anisotropy. So one hopes for a more general explanation.

**Power-law heuristics.** A general but vague idea is that power-law behavior is due to some kind of statistical self-similarity based on a simple scaling invariance principle. For example, consider the Mullins-Sekerka interface motion law:

$$\Delta w = 0 \text{ in } \Omega(t), \quad w = \kappa \text{ on } \Gamma(t) = \partial\Omega, \quad V = \frac{\partial w^+}{\partial \nu} - \frac{\partial w^-}{\partial \nu} \text{ on } \Gamma(t).$$

Changing scale according to  $x = L\tilde{x}$ ,  $t = T\tilde{t}$ ,  $w = A\tilde{w}$  yields

$$\frac{A}{L^2} \Delta \tilde{w} = 0, \quad A\tilde{w} = \frac{1}{L} \tilde{\kappa}, \quad \frac{L}{T} \tilde{V} = \frac{A}{L} \left( \frac{\partial \tilde{w}^+}{\partial \tilde{\nu}} - \frac{\partial \tilde{w}^-}{\partial \tilde{\nu}} \right).$$

This yields a solution of the original system if  $A = 1/L$  and  $L^3 = T$ . Suppose one observes a system undergoing coarsening and plots a length scale  $l$  vs. time  $t$ , so  $l = f(t)$ . Changing scale via  $\tilde{l} = l/L$ ,  $\tilde{t} = t/T$ , one plots  $\tilde{l} = f(T\tilde{t})/T^{1/3} = \tilde{f}(\tilde{t})$ . If the behavior of the system is *independent of scale* (and this concept is very ill-defined for a system with complex morphology), then we can expect that  $f = \tilde{f}$ , and hence, putting  $\tilde{t} = 1$  we get  $f(T) = T^{1/3} f(1)$ . This produces the power-law behavior  $l = ct^{1/3}$ .

**Method of Kohn and Otto.** Recently, a new and rigorous method for explaining power-law behavior was created by Kohn and Otto [27]. The method promises to be rather robust, as it depends only upon a few gross features of the system being considered. It produces *time-averaged power-law bounds* on the decay of a normalized energy  $E(t)$  of the system. These bounds are:

- universal—they apply to every solution.
- one-sided—slower coarsening is possible (the solution can get ‘stuck’ at unstable equilibria, for example), but faster coarsening is *impossible*.
- independent of system complexity (size; morphology of patterns).
- independent of statistical assumptions about the system.

In many of the problems treated so far by the method [28, 29, 30] the power-law bounds are expected to be *typical*, as suggested by the heuristic scaling arguments. (At this time, an exception is that the  $t^{-1/3}$  bounds achieved for a certain mound formation model with anisotropic surface energy having square symmetry do not correspond to the  $t^{-1/4}$  behavior seen in simulations [35, 51].)

With regard to the last bullet above, it is interesting to note that no sort of statistical self-similarity is presumed. In fact, an experimental indication of the significance of this appears recently in work of Voorhees and co-workers [32]. This group experimentally studied 3D coarsening of domain structures in metallic alloys. They observed  $t^{-1/3}$  decay of surface energy over a long range of times where statistics show that morphology and curvature distributions are evolving in a non-self-similar manner. Thus it appears that power-law energy decay is not necessarily indicative of scale-invariant structural behavior.

#### 4.1. *Basic inequalities*

In these notes we will explain the Kohn-Otto method and apply it to the LSW mean-field model and to the monopole model. The essence of the argument involves the consequence of two inequalities that relate a normalized energy  $E(t)$  and a dual quantity  $L(t)$  that loosely characterizes a “length scale” for the system. (It is worth mentioning, though, that an even simpler approach works for the mound coarsening problem treated by Li and Liu in [30].)

**Lemma 4.1:** *Suppose that for all  $t > 0$ , the functions  $E(t)$  and  $L(t)$  satisfy*

$$EL \geq 1 \quad \text{and} \quad \dot{L}^2 \leq -\dot{E}(t), \tag{4.1}$$

*and suppose  $E$  is strictly decreasing. Then, with  $\alpha_1 = \frac{1}{6}$  and  $\alpha_2 = 16$  we have*

$$\int_0^T E(t)^2 dt \geq \alpha_1 \int_0^T (t^{-1/3})^2 dt \quad \text{for all } T \geq \alpha_2 L_0^3. \tag{4.2}$$

**Proof:** 1. Since  $E$  is strictly decreasing, we may say  $L(t) = l(\varepsilon)$ , where  $\varepsilon = E(t)$ . Then

$$\left(\frac{dL}{dt}\right)^2 = \left(\frac{dl}{d\varepsilon} \frac{dE}{dt}\right)^2 \leq -\frac{dE}{dt} \quad \text{implies} \quad \left(\frac{dl}{d\varepsilon}\right)^2 \left(-\frac{dE}{dt}\right) \geq 1.$$

Multiplying by  $E(t)^2$  and integrating, we get

$$f(T) := \int_0^T E(t)^2 dt \geq \int_0^T E^2 \left(\frac{dl}{d\varepsilon}\right)^2 \left(-\frac{dE}{dt}\right) dt = \int_{E_T}^{E_0} \varepsilon^2 \left(\frac{dl}{d\varepsilon}\right)^2 d\varepsilon.$$

2. Next, we fix  $T$  and *minimize* over  $l(\varepsilon)$ : Write  $l = \hat{l} + \tilde{l}$ , where

$$\tilde{l} = 0 \quad \text{at } \varepsilon = E_T \text{ and } E_0, \quad \hat{l} = l = L_0 \text{ at } E_0, L_T \text{ at } E_T.$$

Then

$$\int_{E_T}^{E_0} \varepsilon^2 (\partial_\varepsilon(\hat{l} + \tilde{l}))^2 d\varepsilon = \int_{E_T}^{E_0} \varepsilon^2 ((\partial_\varepsilon \hat{l})^2 + (\partial_\varepsilon \tilde{l})^2) - 2\tilde{l} \partial_\varepsilon(\varepsilon^2 \partial_\varepsilon \hat{l}) d\varepsilon.$$

Choose  $\hat{l}$  such that  $\partial_\varepsilon(\varepsilon^2 \partial_\varepsilon \hat{l}) = 0$ , requiring  $\varepsilon^2 \partial_\varepsilon \hat{l} = \hat{C}$  constant ( $< 0$ ). Then

$$L_0 - L_T = \int_{E_T}^{E_0} \frac{dl}{d\varepsilon} d\varepsilon = \hat{C} \int_{E_T}^{E_0} \varepsilon^{-2} d\varepsilon = \hat{C}(E_T^{-1} - E_0^{-1}).$$

Now it follows

$$f(T) \geq \int_{E_T}^{E_0} \hat{C} \left(\frac{dl}{d\varepsilon}\right) d\varepsilon = \hat{C}(L_0 - L_T) = \frac{(L_T - L_0)^2}{(E_T^{-1} - E_0^{-1})},$$

consequently

$$f(T) \geq E_T(L_T - L_0)^2.$$

3. Next, observe that

$$f'(T) = E_T^2 \geq \frac{1}{L_T^2}, \quad f'(T)f(T)^2 \geq E_T^4(L_T - L_0)^4 \geq \left(1 - \frac{L_0}{L_T}\right)^4.$$

The value  $L_T$  is either large or small:

- If  $L_T \geq 2L_0$ , then  $f' f^2(T) \geq (\frac{1}{2})^4 = \frac{1}{16}$ .
- If  $L_T \leq 2L_0$ , then  $f'(T)L_0^2 \geq \frac{1}{4}$ .

Hence for all  $t$ ,

$$\frac{d}{dt} \left( \frac{f^3}{3} + \frac{fL_0^2}{4} \right) \geq \frac{1}{16},$$

and thus

$$\frac{f(T)^3}{3} + \frac{f(T)L_0^2}{4} \geq \frac{T}{16}.$$

4. We finish as follows. Let  $f = L_0F$ , then  $L_0^3(\frac{1}{3}F^3 + \frac{1}{4}F) \geq \frac{1}{16}T$ . Hence, if  $T \geq 16L_0^3$  it follows  $F^2 \geq \frac{3}{2}$ , since  $F^2 < \frac{3}{2}$  leads to a contradiction:

$$\frac{F^3}{3} + \frac{F}{4} < F \left( \frac{1}{2} + \frac{1}{4} \right) < \sqrt{\frac{3}{2}} \frac{3}{4} < 1 \leq \frac{T}{16L_0^2}.$$

Then if  $T \geq 16L_0^3$  we infer  $\frac{1}{2}F^3 \geq \frac{1}{3}F^3 + \frac{1}{4}F$ , therefore  $\frac{1}{2}f^3 \geq \frac{1}{16}T$ . Thus

$$\int_0^T E(t)^2 dt \geq \frac{T^{1/3}}{2} = \frac{1}{6} \int_0^T (t^{-1/3})^2 dt. \quad \square$$

#### 4.2. *Bounds on coarsening rates for the LSW mean-field model*

Kohn and Otto applied their method to two variants of the Cahn-Hilliard equation in [27]. The method was applied to LSW mean-field models in [15], but Barbara Niethammer has suggested the following argument which exploits the gradient structure of the equations and is very simple.

Suppose we have a collection of spheres of radius  $R_i$  coarsening according to the mean-field law

$$\partial_t R_i = v_i = \frac{1}{R_i} \left( \tilde{\theta} - \frac{2}{R_i} \right).$$

Normalize total volume volume so  $\sum R_i^3 = 1$ . We take

$$E = \sum_i 4\pi R_i^2, \quad L = \sum_i \frac{R_i^4}{4\pi}$$

as energy (total surface area) and “length scale” respectively (normalized by volume). Then by Cauchy-Schwartz we obtain the “interpolation inequality”

$$1 = \left( \sum R_i^3 \right)^2 \leq \left( \sum R_i^2 \right) \left( \sum R_i^4 \right) = EL.$$

The idea to obtain a dissipation inequality is that  $-\dot{E} = g_\Gamma(v, v)$ , coming from the gradient structure of the model. We write

$$-\dot{E} = - \sum 8\pi R_i \dot{R}_i = \sum 4\pi R_i^2 \left( \tilde{\theta} - \frac{2}{R_i} \right) \dot{R}_i = 4\pi \sum R_i^3 \dot{R}_i^2.$$

Then

$$(\dot{L})^2 = \left( \sum \frac{R_i^3 \dot{R}_i}{\pi} \right)^2 \leq \left( \sum R_i^3 \right) \left( \sum R_i^3 \dot{R}_i^2 \right) \leq -\dot{E}.$$

Then we simply apply the Lemma to obtain time-averaged bounds on the coarsening rate, universally valid for any size distribution of spheres:

$$\int_0^T E(t)^2 dt \geq \frac{1}{6} \int_0^T (t^{-1/3})^2 dt \quad \text{for } T \geq 16L_0^2.$$

The same calculation can be made for the Lifshitz–Slyozov conservation law:

$$\partial_t n(R, t) + \partial_R(Vn) = 0, \quad V(R, t) = \frac{1}{R} \left( \tilde{\theta} - \frac{2}{R} \right),$$

using the *characteristic map*  $\hat{R}(r, t)$  satisfying

$$\partial_t \hat{R} = V(\hat{R}, t), \quad \hat{R}(r, 0) = r,$$

and the fact that  $n(R, t) dR = n_0(r) dr$  with  $R = \hat{R}(r, t)$ . We omit the details.

### 4.3. Bounds on coarsening rates for the monopole model

As discussed above, the monopole model is important because of its status as a model of multi-dimensional coarsening for which extensive computer simulations are feasible and have been carried out. The spatial distribution of particle positions affects coarsening dynamics in ways that are not well understood; presumably correlations develop between neighboring particles in a time-dependent fashion. As we now show, however, it is not difficult to establish a basic power-law bound on the coarsening rate using the Kohn–Otto method.

We recall that  $\dot{R}_i = v_i$ , where

$$R_i^3 v_i + \sum_{i \neq j} \frac{R_i^2 R_j^2 v_j}{|x_i - x_j|} = R_i^2 \left( \tilde{\theta} - \frac{2}{R_i} \right), \quad \sum R_i^2 v_i = 0.$$

Note that the field  $u = \mathcal{T}v$  satisfies

$$\Delta u = 0 \text{ in } \mathbb{R} \setminus \Gamma, \quad [n \cdot \nabla u]_{-}^{+} = v \text{ on } \Gamma, \quad u \rightarrow 0 \text{ as } |x| \rightarrow \infty, \quad (4.3)$$

and that

$$\int_{\Gamma} (u + \kappa) \tilde{v} = 0 \quad \text{for all } \tilde{v} \text{ with } \int_{\Gamma} \tilde{v} = 0.$$

Again, the surface area (surface energy)  $E = \sum_i 4\pi R_i^2 = \int_{\Gamma} 1$  satisfies

$$\begin{aligned} -\dot{E} &= -\sum 8\pi R_i \dot{R}_i = \sum 4\pi R_i^2 \left( -\frac{2}{R_i} v_i \right) \\ &= \int_{\Gamma} -\kappa v = \int_{\Gamma} uv = \int_{\mathbb{R}^3} |\nabla u|^2 = g_{\Gamma}(v, v). \end{aligned}$$

We use a rather fancy “length scale”  $L$  determined by

$$L^2 = \int_{\mathbb{R}^3} |\nabla \Psi|^2, \quad \text{where } -\Delta \Psi = \mathbf{1}_{\cup B_i},$$

meaning  $-\Delta \Psi = 1$  on the union of balls  $\cup_i B_i$  and 0 on the complement. The function  $\Psi$  is a superposition  $\Psi = \sum \Psi_i$ , where by scaling,

$$\Psi_i(x) = \phi \left( \frac{|x - x_i|}{R_i} \right) R_i^2.$$

The function  $\phi$  satisfies  $-(\partial_r^2 + \frac{2}{r}\partial_r)\phi = \mathbf{1}_{r < 1}$ , and is given by

$$\phi(r) = \frac{1}{2} - \frac{1}{6}r^2 \text{ for } r < 1, \quad \frac{1}{3r} \text{ for } r > 1.$$

(Note  $\partial_r \phi = -1/3$  at  $r = 1^{\pm}$ .) Thus

$$\Psi_i = \frac{R_i^2}{2} - \frac{r^2}{6} \text{ for } r = |x - x_i| < R_i, \quad \frac{R_i^3}{3r} \text{ for } r > R_i.$$

We notice that

$$\dot{\Psi}_i = R_i \dot{R}_i \text{ for } r < R_i, \quad \frac{R_i^2 \dot{R}_i}{r} \text{ for } r > R_i,$$

hence

$$\Delta \dot{\Psi} = 0 \text{ in } \mathbb{R} \setminus \Gamma, \quad [n \cdot \nabla \dot{\Psi}]_{-}^{+} = -\dot{R}_i \text{ on } \Gamma_i, \quad \dot{\Psi}(x) \rightarrow 0 \text{ as } |x| \rightarrow \infty.$$

Comparing with (4.3) we see that  $\dot{\Psi} = u = \mathcal{T}v$ . By consequence, we get the *dissipation inequality* as follows:

$$(L\dot{L})^2 = \left( \int_{\mathbb{R}^3} \nabla \Psi \cdot \nabla \dot{\Psi} \right)^2 \leq \int_{\mathbb{R}^3} |\nabla \Psi|^2 \int_{\mathbb{R}^3} |\nabla \dot{\Psi}|^2 = L^2 \int_{\mathbb{R}^3} |\nabla u|^2 = L^2(-\dot{E}).$$

Hence  $(\dot{L})^2 \leq -\dot{E}$ .

We prove the *interpolation inequality*  $EL \geq 1$  by using an explicit expression for  $L^2$  obtained as follows:

$$L^2 = \int_{\mathbb{R}^3} \Psi(-\Delta\Psi) = \int_{\cup B_i} \Psi = \sum_{i,j} \int_{B_i} \Psi_j$$

We evaluate the integrals using scaling for  $i = j$  and the mean value property for the harmonic function  $\Psi_j$  on  $B_i$  for  $i \neq j$ :

$$\int_{B_i} \Psi_i = R_i^{2+3} \int_0^1 \left(\frac{1}{2} - \frac{r^2}{6}\right) 4\pi r^2 dr = R_i^5 \frac{8\pi}{15},$$

$$\int_{B_i} \Psi_j = \frac{4\pi}{3} R_i^3 \Psi_j(x_i) = \frac{4\pi}{3} \frac{R_i^3 R_j^3}{3|x_i - x_j|}.$$

Hence

$$L^2 = \sum_i \frac{8\pi}{15} R_i^5 + \sum_{j \neq i} \frac{4\pi}{3} \frac{R_i^3 R_j^3}{3|x_i - x_j|} \geq \sum_i \frac{8\pi}{15} R_i^5. \tag{4.4}$$

Normalizing so that  $\sum R_i^3 = 1$ , we get

$$1 = \sum R_i^3 = \left(\sum R_i^{\frac{4}{3}} R_i^{\frac{5}{3}}\right)^{\frac{3}{2}} \leq \left(\sum R_i^2\right) \left(\sum R_i^5\right)^{\frac{1}{2}}$$

by Hölder’s inequality with  $p = \frac{3}{2}$ ,  $q = 3$  ( $p^{-1} + q^{-1} = 1$ ). Hence  $EL \geq 1$ .

Now applying the ODE lemma yields

$$\int_0^T E(t)^2 dt \geq \frac{1}{6} \int_0^T (t^{-1/3})^2 dt \quad \text{for all } T \geq 16L_0^3.$$

This estimate provides a power-law bound on energy dissipation with the exponent one expects physically. We remark, however, that the inequality (4.4) is likely to be quite pessimistic in dense systems, and so one might hope for improvements regarding the prefactor or time of validity.

## 5. Smoluchowski's coagulation equations

### 5.1. Introduction

A simple mean-field model of coarsening by coalescence leads to Smoluchowski's coagulation equations. Particles or clusters of size  $x$  and size  $y$  combine at a rate proportional to the population of each and a rate kernel  $K(x, y)$ . Schematically, the number of particles of size  $x$  is affected by the processes



One writes a mean-field rate equation for the number density  $n(x, t)$  in the form

$$\partial_t n(x, t) = \frac{1}{2} \int_0^x K(x-y, y) n(x-y, t) n(y, t) dy - \int_0^\infty K(x, y) n(x, t) n(y, t) dy. \quad (5.1)$$

Perhaps due to its simplicity, this model has found an amazingly diverse set of applications over a vast range of scales. It has been used to study microdroplet formation (in clouds, ink fog, smoke, fuel, paint, etc.), the kinetics of polymerization, hashing algorithms, and the clustering of colloids, phytoplankton in "marine snow," planetesimals in stellar accretion disks, and stars themselves. Much scientific effort has gone into determining appropriate rate kernels  $K(x, y)$  for different physical models. For simplicity, here we will only consider the constant kernel  $K = 2$ , for which we can get a solution formula via the Laplace transform.

Since particles only combine, one expects the size distribution to shift toward larger particles and the typical particle size to grow in time. So one must rescale size to observe nontrivial long-time dynamical behavior. Equation (5.1) has the scaling invariance that if  $n(x, t)$  is a solution, then so is

$$\tilde{n}(x, t) = a n(bx, ct)$$

provided  $a = bc$ .

### 5.2. A 'new' framework for dynamic scaling analysis

An issue of considerable significance in applications is whether typically the size distribution will approach a scale-invariant form, for which  $\tilde{n} = n$ . This is a self-similar solution or *scaling* solution for short.

In these notes we will study this question within a larger framework for understanding dynamic scaling behavior that has been outlined in the paper [33]. The basic issues are:

- (i) What scaling solutions exist?
- (ii) What are their domains of attraction?
- (iii) What are the most general scaling limit points?  
(These comprise the “scaling attractor.”)
- (iv) How can we describe the dynamics on the set of limit points?  
(This is arguably the “ultimate dynamics” of the system.)
- (v) How complicated can the ultimate dynamics be?

Though evidently stated in dynamical terms, this framework is strongly motivated by classical results in probability theory that date back to the 1920s, involved with establishing necessary and sufficient conditions for convergence in the central limit theorem. The issue concerns scaling limits of sums  $S_n = \sum_{j=1}^n X_j$  of independent and identically distributed random variables, as  $n \rightarrow \infty$ . The whole theory is beautifully exposed in Feller’s great book [19], and provides complete answers to the questions in the framework above:

- The normal distribution is the unique scale-invariant distribution of finite variance. But more generally the class of scale-invariant distributions make up a two-parameter family of (heavy-tailed) distributions called the Lévy stable laws.
- The normal distribution attracts all distributions of finite variance. But in general, simple necessary and sufficient conditions for a scaling limit to exist are known in terms of the power-law behavior (*regular variation*, to be precise) of the second-moment distribution function.
- The most general scaling limits that can arise for some subsequence  $n_j \rightarrow \infty$  are the *infinitely divisible distributions*. These form an infinite-dimensional family parametrized by the famous Lévy-Khintchine representation formula in terms of a measure satisfying certain finiteness conditions.
- There exist distributions (Doebelin’s universal laws) for which *every possible* scaling limit is realized along some subsequence. This is a hallmark of chaos—sensitive dependence on initial conditions.

The analytical methods used to establish these classical limit theorems are not essentially probabilistic in nature. Rather, they appear to be natural tools for analyzing scaling dynamics in many problems. For Smoluchowski’s equation with  $K = 2$ , the results that we will (mostly) prove here are stated informally as follows (see [34, 33]):

- We find a unique scale-invariant solution with finite number  $\int_0^\infty n \, dx$  and mass  $\int_0^\infty xn \, dx$ . But there is a one-parameter family of infinite-mass self-similar solutions, with profiles given by *Mittag-Leffler* probability distribution functions.
- Solutions approach self-similar form as  $t \rightarrow \infty$  if and only if initially  $x \mapsto \int_0^x yn \, dy$  is almost power-law—regularly varying at  $\infty$ .
- The scaling attractor (set of subsequential scaling limits) can be parametrized by measures satisfying certain finiteness conditions.
- The nonlinear dynamics on the scaling attractor is *linearized* in terms of this measure representation.
- This ultimate dynamics exhibits sensitive dependence on initial data.

These results are strikingly analogous to those of classical probability theory. For well-localized data (finite mass), there is one universal scaling behavior, analogous to the central limit theorem. For many physical applications this is the most relevant case. However, one can study scaling dynamics in a more general context, and there one finds a rich set of mathematical possibilities. These should not be dismissed as uninteresting, given the wide range of applications of Smoluchowski's model. Heavy-tailed distributions have come to be important in numerous applications of probability, for example.

### 5.3. Solution by Laplace transform

The rigorous study of solutions of the coagulation equation (5.1) begins with the general moment identity

$$\begin{aligned} \partial_t \int_0^\infty a(x)n(x,t) \, dx &= \frac{1}{2} \int_0^\infty \int_0^x a(x)n(x-y,t)n(y,t)K(x-y,y) \, dy \, dx \\ &\quad - \int_0^\infty \int_0^\infty a(x)n(x,t)n(y,t)K(x,y) \, dy \, dx \\ &= \frac{1}{2} \int_0^\infty (a(x+y) - a(x) - a(y))n(x,t)n(y,t)K(x,y) \, dx \, dy. \end{aligned} \quad (5.2)$$

Taking  $a = x$  formally yields conservation of total mass:

$$\partial_t \int_0^\infty xn(x,t) \, dx = 0.$$

For  $K = 2$ , the total number  $N(t) = \int_0^\infty n(x,t) \, dx$  satisfies  $\partial_t N = -N^2$ . We will normalize and scale  $x$  and  $t$  so that we always have

$$N(t) = \frac{1}{t}. \quad (5.3)$$

Next take  $a(x) = 1 - e^{-qx}$ . Since  $a(x + y) - a(x) - a(y) = -a(x)a(y)$ , with

$$\phi(q, t) = \int_0^\infty (1 - e^{-qx})n(x, t)dt, \tag{5.4}$$

a quantity related to the Laplace transform of  $n$ , we have the simple equation

$$\partial_t \phi = -\phi^2. \tag{5.5}$$

Note that

$$\phi(0, t) = 0, \quad \phi(\infty, t) = N(t) = \frac{1}{t}, \quad \partial_q \phi = \int_0^\infty e^{-qx} xn(x, t) dx. \tag{5.6}$$

Since  $\partial_t(1/\phi) = 1$ , for  $t, t_0 > 0$  we obtain the solution formulae

$$\frac{1}{\phi(q, t)} - \frac{1}{\phi(q, t_0)} = t - t_0, \quad \phi(q, t) = \frac{\phi(q, t_0)}{1 + (t - t_0)\phi(q, t_0)}. \tag{5.7}$$

This solution formula serves as the basis for a theory of the initial-value problem for which the size distribution  $n(x, t) dx = \nu_t(dx)$  is a finite measure on  $(0, \infty)$  which initially can be completely arbitrary, subject to the normalization

$$\int_0^\infty \nu_{t_0}(dx) = \frac{1}{t_0}. \tag{5.8}$$

Let us sketch how this works. First, look for solutions that are lattice measures, of the form

$$\nu_t(dx) = \sum_{j=1}^\infty c_j(t)\delta(x - j\Delta x)$$

where  $\delta(x - j\Delta x)$  is a Dirac mass at  $j\Delta x$ . With initially  $\sum_{j=1}^\infty c_j(t_0) = 1/t_0$ , solve the discrete equations

$$\partial_t c_j = \sum_{k=1}^{j-1} c_{j-k}c_k\Delta x - 2c_jN(t) \tag{5.9}$$

inductively for  $j = 1, 2, \dots$ , with  $N(t) = 1/t$ . Then prove that

$$\sum_{j=1}^\infty c_j(t) \leq \frac{1}{t} \quad \text{for all } t \geq t_0.$$

(Hint:  $N_J(t) = \sum_{j=1}^J c_j(t)$  satisfies  $\partial_t(N_J - N) \leq (N_J - N)^2$ ,  $N_J(t_0) - N(t_0) \leq 0$ .) Integrate (5.9) in  $t$  and apply the Laplace transform to deduce that the function  $\phi(q, t) = \int_0^\infty (1 - e^{-qx})\nu_t(dx)$  satisfies

$$\phi(q, t) = \phi(q, t_0) + \int_{t_0}^t \phi(q, s)^2 ds.$$

This implies (5.7), and we infer that  $t \mapsto \nu_t$  is weakly continuous in the sense of measures by the continuity theorem for Laplace transforms. In general, we approximate a general measure  $\nu_{t_0}$  by lattice measures as above. We pass to limits using the continuity theorem for Laplace transforms together with (5.7):

$$\begin{aligned} \nu_{t_0}^{\Delta x} \rightarrow \nu_{t_0} &\Leftrightarrow \phi_0^{\Delta x}(q) \rightarrow \phi(q, t_0) \text{ for all } q > 0 \\ &\Leftrightarrow \phi^{\Delta x}(q, t) \rightarrow \phi(q, t) \text{ for all } q \geq 0, \quad t \geq t_0 \end{aligned}$$

We obtain existence and uniqueness, and solutions depend continuously on initial data with respect to weak convergence. Also, initial data depend continuously on the solution! See [34] for discussion of a precise sense in which this yields a weak solution of Smoluchowski’s equation. The upshot is that for any measure  $\nu_{t_0}$  on  $(0, \infty)$  satisfying (5.8) with  $t_0 > 0$ , there is a unique measure solution defined for all  $t \geq t_0$ , meaning a weakly continuous map  $t \mapsto \nu_t$  such that for  $t \geq t_0$ ,  $\nu_t$  is a finite measure on  $(0, \infty)$  such that  $\phi(q, t) = \int_0^\infty (1 - e^{-qx})\nu_t(dx)$  satisfies (5.7).

#### 5.4. Scaling solutions and domains of attraction

Based upon the solution of the coagulation equation by Laplace transform, a complete classification of scaling solutions and their domains of attraction was worked out in [34] for  $K = 2$ ,  $x + y$  and  $xy$ . For  $K = 2$ , all nontrivial scaling limits can be classified as follows.

**Theorem 5.1:** *Take  $t_0 = 1$  and suppose  $\nu_t$  is a measure solution of Smoluchowski’s equation, so that (5.7) holds, and introduce the probability distribution function*

$$F_t(x) = \int_0^x \nu_t(dx) / \int_0^\infty \nu_t(dx) \quad (= t \int_0^x n(y, t) dy).$$

(i) *Suppose that there exists  $\lambda(t) \rightarrow \infty$  and a probability distribution  $F_*$  so that*

$$F_t(\lambda(t)x) \rightarrow F_*(x) \quad \text{as } t \rightarrow \infty \tag{5.10}$$

*at all points of continuity, where  $F_*(x) < 1$  for some  $x > 0$ . Then,*

$$\int_0^x y\nu_1(dy) \sim x^{1-\rho}L(x) \quad \text{as } x \rightarrow \infty \tag{5.11}$$

for some  $\rho \in (0, 1]$  and  $L$  slowly varying at  $\infty$ .

(ii) Conversely, suppose that (5.11) holds. Then (5.10) holds, with

$$F_* = F_\rho(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} x^{\rho k}}{\Gamma(1 + \rho k)},$$

a Mittag-Leffler distribution, whose Laplace transform is

$$\mathcal{L}F_\rho(q) = \int_0^\infty e^{-qx} F_\rho(dx) = \frac{1}{1 + q^\rho}.$$

**Remark 5.2:** Finite mass

$$\int_0^\infty x\nu_1(dx) < \infty$$

gives  $\rho = 1$ , and  $F_1(x) = 1 - e^{-x}$  corresponding to

$$n(t, x) = \frac{1}{t^2} e^{-x/t}.$$

This is an analog of the central limit theorem in probability theory. The Mittag-Leffler distributions  $F_\rho$  for  $0 < \rho < 1$  have infinite mass and are analogs of the (heavy-tailed) Lévy stable laws of probability theory.

**Proof:** The strategy of the proof is to use the rigidity property of scaling limits, and the Tauberian theorem (mentioned in the Background section above). Assume, as in the statement of the theorem, that there exists  $\lambda(t) \rightarrow \infty$  and a nontrivial probability distribution  $F_*$  so that

$$F_t(\lambda(t)x) \rightarrow F_*(x) \quad \text{as } t \rightarrow \infty,$$

at all points of continuity. By the continuity theorem for Laplace transforms,

$$\int_0^\infty e^{-qx} F_t(\lambda(x) dx) \rightarrow \int_0^\infty e^{-qx} F_*(dx) = \mathcal{L}F_*(q) \quad \text{for all } q > 0.$$

In terms of  $\phi$  and  $\phi_1(q) := \phi(q, 1)$ , this means that for all  $q > 0$ ,

$$1 - t\phi(q/\lambda, t) = 1 - \frac{t\phi_1(q/\lambda)}{1 + (t - 1)\phi_1(q/\lambda)} \rightarrow \mathcal{L}F_*(q) \in (0, 1).$$

Therefore, for all  $q > 0$ ,

$$t\phi_1(q/\lambda(t)) \rightarrow g(q) \in (0, \infty) \quad \text{as } t \rightarrow \infty.$$

By the rigidity property for scaling limits, we must have  $g(q) = cq^\rho$  for some  $c > 0$ ,  $0 \leq \rho < \infty$ . Then

$$\mathcal{L}F_*(q) = 1 - \frac{cq^\rho}{1 + cq^\rho} = \frac{1}{1 + cq^\rho}.$$

Since  $F_*(\infty) = \mathcal{L}F_*(0+) = 1$ , we have  $\rho > 0$ . Since  $-\partial_q \mathcal{L}F_* = \int_0^\infty e^{-qx} x F_*(dx)$  is positive decreasing, we have  $\rho \leq 1$ .

By scaling  $\lambda(t)$ , we can achieve  $c = 1$ . Moreover, the rigidity property implies

$$\phi_1(q) \sim q^\rho \tilde{L}(1/q) \quad \text{as } q \rightarrow 0^+,$$

for some  $\tilde{L}$  slowly varying at  $\infty$ . By the Tauberian theorem, the result

$$\int_0^x y \nu_1(dy) \sim x^{1-\rho} L(x) \quad \text{as } x \rightarrow \infty$$

is equivalent to

$$\partial_q \phi_1(q) = \int_0^\infty e^{-qy} y \nu_1(dy) \sim q^{\rho-1} L(1/q) \Gamma(2-\rho), \quad \text{as } q \rightarrow 0^+.$$

The proof (in both directions) is finished with the use of the following lemma. □

**Lemma 5.3:** *The following are equivalent.*

- (1)  $\phi_1(q) \sim q^\rho \tilde{L}(1/q)$  as  $q \rightarrow 0^+$ .
- (2)  $\phi'_1(q) \sim \rho q^{\rho-1} \tilde{L}(1/q)$  as  $q \rightarrow 0^+$ .

**Proof:** We first show that (1)  $\Rightarrow$  (2). Since  $\phi''_1 = -\int_0^\infty e^{-qy} y^2 \nu_1(dy) < 0$ , we have that for fixed  $a > 1$ ,

$$\phi'_1(q) \geq \frac{\phi_1(aq)\phi_1(q)}{aq - q} = \frac{(aq)^\rho \hat{L}(aq) - q^\rho \hat{L}(q)}{q(a-1)} = q^{\rho-1} \hat{L}(q) \frac{a^\rho \hat{L}(aq)/\hat{L}(q) - 1}{a-1}$$

and  $\hat{L}(q) = \phi_1(q)/q \sim \tilde{L}(1/q)$ . Hence (take  $a \rightarrow 1$ )

$$\liminf_{q \rightarrow 0} \frac{\phi'_1(q)}{q^{\rho-1} \hat{L}(q)} \geq \frac{a^\rho - 1}{a - 1} \rightarrow \rho.$$

Similarly, for fixed  $a < 1$ , we get

$$\limsup_{q \rightarrow 0} \frac{\phi'_1(q)}{q^{\rho-1} \hat{L}(q)} \leq \frac{a^\rho - 1}{a - 1} \rightarrow \rho.$$

In order to show that (2)  $\Rightarrow$  (1), let  $\hat{L}(q) = \phi'_1(q)/\rho q^{\rho-1} \sim \tilde{L}(1/q)$ . Then

$$\phi_1(q) = \int_0^q \phi'_1(s) ds = \int_0^q \rho s^{\rho-1} \hat{L}(s) ds = q^\rho \hat{L}(q) \int_0^1 \rho x^{\rho-1} \frac{\hat{L}(qx)}{\hat{L}(q)} dx,$$

where in the last equality, we changed variables according to  $s = qx$ . We know

$$\frac{\hat{L}(qx)}{\hat{L}(q)} \rightarrow 1, \quad \text{for all } x > 0,$$

and we also need the fact (not difficult to prove) that  $\tilde{L}(x) \geq C_\varepsilon x^{-\varepsilon}$  for all  $\varepsilon > 0$ . For fixed  $\varepsilon$  we use dominated convergence to conclude that

$$\frac{\phi_1(q)}{q^\rho \hat{L}(q)} \rightarrow \int_0^1 \rho x^{\rho-1} dx = 1, \quad \text{as } q \rightarrow 0. \quad \square$$

### 5.5. The scaling attractor

In systems with complicated dynamics, a fundamental notion aimed at capturing all long-time behavior, not only limiting states as  $t \rightarrow \infty$ , is that of the attractor. In finite-dimensional systems, one definition describes the attractor in terms of all possible limit points of bounded sequences of solutions. Modulo rescaling in size, this is precisely what we aim to describe here, following [33]. The resulting object, which we call the *scaling attractor*, turns out to have a remarkable characterization analogous to the Levy-Khinchine representation of infinitely divisible laws in probability theory.

**Definition 5.4:** Suppose  $\hat{F}$  is a probability distribution function such that there exists a sequence of solutions  $\nu_t^{(n)}$  defined for  $t \geq t_0$  and numbers  $t_n, \beta_n \rightarrow \infty$  such that

$$F_{t_n}^{(n)}(\beta_n x) \rightarrow \hat{F}(x) \quad \text{as } n \rightarrow \infty$$

at each point of continuity. Then we say that  $\hat{F}$  belongs to the (*proper*) *scaling attractor*  $\mathcal{A}$ .

One property enjoyed by the attractor in a finite system is that it is an invariant set forward and backward in time. A related property holds for the scaling attractor. First we note the following scaling property for measure solutions of Smoluchowski's equation with  $K = 2$ : Let  $a > 0, b > 0$  be given, and let  $\nu_t$  be a solution on  $[t_0, \infty)$ . Now, let

$$\tilde{\nu}_t(dx) = a\nu_{at}(b dx),$$

with  $\tilde{F}_t(x) = F_{at}(bx)$ . Then,  $\tilde{\nu}$  is again a solution, on  $[t_0/a, \infty)$ , because of the fact that

$$\tilde{\phi}(q, t) = \int_0^\infty (1 - e^{-qx})\nu_{at}(b dx) = a\phi(q/b, at)$$

satisfies  $\partial_t \tilde{\phi} = a^2 (\partial_t \phi)(q/b, at) = -a^2 \phi^2 = -\tilde{\phi}^2$ .

Now, suppose we have a sequence as in the definition above. Put

$$\tilde{F}_t^{(n)}(x) = F_{t_n t}^{(n)}(\beta_n x).$$

Then

$$\tilde{F}_1^{(n)}(x) \rightarrow \hat{F}(x) \text{ as } n \rightarrow \infty.$$

Correspondingly,  $\tilde{\phi}^{(n)}(q, 1) \rightarrow \hat{\phi}(q)$  by the continuity theorem, and

$$\tilde{\phi}^{(n)}(q, t) = \frac{\tilde{\phi}^{(n)}(q, 1)}{1 + (t-1)\tilde{\phi}^{(n)}(q, 1)} \rightarrow \frac{\hat{\phi}(q)}{1 + (t-1)\hat{\phi}(q)} =: \phi(q, t).$$

We have

$$\phi(q, t) = \int_0^\infty (1 - e^{-qx}) \nu_t(dx),$$

where  $\nu_t$  is a solution on  $[t_1, \infty)$  for all  $t_1 > 0$ . Starting from any  $t_0 > 0$ , such a solution is defined *backwards in time* as far as it is meaningful.

**Definition 5.5:** A solution with  $K = 2$  defined for all  $t > 0$  is an *eternal solution*.

This analysis proves the following:

**Theorem 5.6:** *Points in the scaling attractor correspond one-to-one with eternal solutions. That is,  $\hat{F} \in \mathcal{A}$  if and only if  $\hat{F} = F_1$  for some eternal solution  $\nu_t$ .*

We get an interesting characterization of the scaling attractor by studying limits as  $t \downarrow 0$ . Observe that

$$\phi(q, t) = \frac{\hat{\phi}(q)}{1 + (t-1)\hat{\phi}(q)} \rightarrow \frac{\hat{\phi}(q)}{1 - \hat{\phi}(q)} =: \Phi(q), \text{ as } t \rightarrow 0^+.$$

This raises the question: What does this mean in terms of weak convergence of measures? Note that  $\int_0^\infty \nu_t(dx) = 1/t \rightarrow \infty$  as  $t \rightarrow \infty$ , and  $\hat{\phi}(q) \rightarrow 1$  as  $q \rightarrow 0^+$ , so  $\Phi(\infty) = \infty$ . Also  $t\phi(q, t) \rightarrow 0$ , so the limit  $F_t \rightarrow \delta(x-0)$  is trivial.

The answer turns out to be to look at  $G_t(dx) = x\nu_t(dx)$ . One has

$$\phi(q, t) = \int_0^\infty \frac{1 - e^{-qx}}{x} G_t(dx), \quad \partial_q \phi(q, t) = \int_0^\infty e^{-qx} G_t(dx) = \mathcal{L}G_t(q).$$

Since

$$\frac{1}{\phi} = \frac{1}{\hat{\phi}} + (t-1),$$

we get that  $\partial_q \phi / \phi^2 = \partial_q \hat{\phi} / \hat{\phi}(q)^2$ , and hence as  $t \rightarrow 0$ ,

$$\partial_q \phi(q, t) = \mathcal{L}G_t(q) \rightarrow \frac{\Phi^2}{\hat{\phi}^2} = \frac{\partial_q \hat{\phi}}{(1 - \hat{\phi})^2} = \partial_q \Phi(q).$$

By the extended continuity theorem, there exists a measure  $H$  on  $[0, \infty)$  with

$$G_t \rightarrow H \quad \text{as } t \rightarrow 0^+.$$

From

$$\phi(q, t) = \phi(\varepsilon, t) + \int_\varepsilon^q \partial_q \phi(q', t) dq',$$

taking  $t \rightarrow 0$  and  $\varepsilon \downarrow 0$ , we get

$$\Phi(q) = \int_0^q \mathcal{L}H(q') dq' = \int_0^\infty \frac{1 - e^{-qx}}{x} H(dx).$$

Since  $\Phi(\infty) = \infty$ ,

$$\text{either } H(0) > 0 \quad \text{or} \quad \int_0^\infty x^{-1} H(dx) = \infty. \tag{5.12}$$

**Definition 5.7:** A measure  $G$  on  $[0, \infty)$  is a *generating measure* if

$$\int_{[0,x]} G(dy) + \int_{[x,\infty)} y^{-1} G(dy) < \infty \quad \text{for all } x > 0,$$

i.e.,

$$\int_{[0,\infty)} (1 \wedge x^{-1}) G(dx) < \infty.$$

$G$  is *divergent* if either  $G(0) > 0$  or  $\int_{(0,\infty)} y^{-1} G(dy) = +\infty$ .

**Theorem 5.8:** To each non-divergent generating measure  $\hat{G}$  with

$$\int_{(0,\infty)} y^{-1} \hat{G}(dy) = \frac{1}{t_0}, \quad t_0 > 0,$$

there corresponds a unique solution  $\nu_t$  on  $[t_0, \infty)$  with  $\nu_{t_0}(dx) = x^{-1} \hat{G}(dx)$ , and conversely. Furthermore, to each eternal solution  $\nu_t$  on  $(0, \infty)$  there corresponds a divergent generating measure  $H$ , such that

$$G_t \rightarrow H \quad \text{as } t \rightarrow 0.$$

Conversely, to each divergent generating measure  $H$  corresponds a unique eternal solution  $\nu_t$  as above, determined by

$$\phi(q, t) = \frac{\Phi(q)}{1 + t\Phi(q)}, \quad \Phi(q) = \int_{[0,\infty)} \frac{1 - e^{-qx}}{x} H(dx).$$

**Proof:** We will prove the converse result. For small  $\varepsilon > 0$  put

$$\hat{G}_\varepsilon = H|_{[\varepsilon, \infty)} + H(0)\delta(x - \varepsilon).$$

Define  $t_0(\varepsilon)$  by

$$\frac{1}{t_0(\varepsilon)} = \int_{(0, \infty)} x^{-1} \hat{G}_\varepsilon(dx) = \int_{[\varepsilon, \infty)} x^{-1} H(dx) + \frac{1}{\varepsilon} H(0).$$

Then  $t_0(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ , and  $\hat{G}_\varepsilon \rightarrow H$  on  $[0, \infty)$ .  $\hat{G}_\varepsilon$  determines a solution  $\nu_t^\varepsilon$  on  $[t_0(\varepsilon), \infty)$  with

$$\phi_0^\varepsilon(q) = \phi^\varepsilon(q, t_0(\varepsilon)) = \int_0^\infty \frac{1 - e^{-qx}}{x} \hat{G}_\varepsilon(dx) \rightarrow \int_0^\infty \frac{1 - e^{-qx}}{x} H(dx) = \Phi(q),$$

as  $\varepsilon \rightarrow 0$ . Then, for  $t > t_0(\varepsilon)$ ,

$$\phi^\varepsilon(q, t) = \frac{\phi_0^\varepsilon(q)}{1 + (t - t_0(\varepsilon))\phi_0^\varepsilon(q)} \rightarrow \frac{\Phi(q)}{1 + t\Phi(q)} =: \phi(q, t),$$

corresponding to a unique eternal solution  $\nu_t$ . □

The result of this theorem is that *each point on the scaling attractor  $\mathcal{A}$  corresponds to a unique divergent generating measure  $H$ :*

$$F_1 \in \mathcal{A} \longleftrightarrow \nu_t \text{ is eternal} \longleftrightarrow H \text{ is a divergent generating measure.}$$

### 5.6. Linearization of dynamics on the scaling attractor

Understanding the dynamics on the scaling attractor in terms of the measures  $H$  turns out to be a simple consequence of a scaling property of solutions. The upshot is that nonlinear clustering dynamics governed by the coagulation equation (5.1) with  $K = 2$  becomes *linear* in terms of  $H$ .

Suppose  $\nu_t$  is an eternal solution. Given  $a, b > 0$ , let  $\tilde{\nu}_t(dx) = a\nu_{at}(b dx)$  and  $\tilde{F}_t(x) = F_{at}(bx)$ . Then, since  $\nu_t$  is eternal,  $\tilde{\nu}_t$  is also eternal and furthermore,  $\tilde{F}_1(x) = F_a(bx)$ . Also, we have

$$\tilde{G}_t \rightarrow \tilde{H} \quad \text{weakly as } t \rightarrow 0.$$

But

$$\tilde{G}_t(dx) = x\tilde{\nu}_t(dx) = ax\nu_{at}(b dx) = \frac{a}{b}G_{at}(b dx) \rightarrow \frac{a}{b}H(b dx), \text{ as } t \rightarrow 0.$$

Hence

$$\tilde{H}(x) = \frac{a}{b}H(bx).$$

**Theorem 5.9:** Under the correspondence  $\mathcal{G}$  that maps the scaling attractor  $\mathcal{A}$  to divergent generating measures, given by

$$\mathcal{G}(F_1) = H_1 = H,$$

the scaling dynamics on  $\mathcal{A}$  given by  $F_1(x) \mapsto F_t(bx)$  is represented via the map

$$H(dx) \mapsto H_t(dx) = \mathcal{G}(F_t(bdx)) = \frac{t}{b} H_1(bdx).$$

**Remark 5.10:** In greatest generality, the scaling dynamics is complicated! One can show that there exists an  $H_1$  so that the trajectory  $t \mapsto F_t$  is *dense* in  $\mathcal{A}$ . This means that solutions exhibit sensitive dependence on initial data, the hallmark of chaos. To show this, basically we need to show that for every divergent generating measure  $H$ , there exist  $t_n, b_n \rightarrow \infty$  such that

$$\frac{t_n}{b_n} H_1(b_n dx) \rightarrow \hat{H},$$

in an appropriate topology. One arranges this by carefully “packing the tail” of the measure  $H_1$  in a way similar to the construction of Doebelin’s universal laws in probability [19]. Details will appear in [33].

**Self-similar solutions.** These correspond to solutions invariant under continuous rescaling with  $b(t) \rightarrow \infty$  as  $t \rightarrow \infty$ , so that  $F_t(b(t)x) = F_1(x)$ , i.e.,

$$H(x) = \frac{t}{b(t)} H(b(t)x) \quad \text{for all } x > 0.$$

Take  $t \rightarrow \infty$  and apply the rigidity lemma. Then  $H$  must be a pure power:

$$H(x) = \tilde{c}x^p \quad \text{for some } \tilde{c} > 0, 0 \leq p < \infty.$$

Then

$$\partial_q \Phi(q) = \int_0^\infty e^{-qx} H(dx) = \tilde{c} \int_0^\infty e^{-qx} d(x^p) = \tilde{c}q^{-p},$$

so  $\Phi(q) = cq^{1-p}$ . Note that  $p < 1$ , since  $\int_1^\infty x^{-1} H(dx) < \infty$ . Then,

$$t\phi(q, t) = \int_0^\infty (q - e^{-qx}) F_t(dx) = \frac{tcq^{1-p}}{1 + tcq^{1-p}}$$

gives

$$\mathcal{L}F_t(q) = 1 - t\phi = \frac{1}{1 + tcq^{1-p}} = \mathcal{L}\hat{F}_1((ct)^{\frac{1}{1-p}} q),$$

with

$$\mathcal{L}\hat{F}_1(q) = \frac{1}{1 + q^{1-p}} = \frac{1}{1 + q^\rho}, \quad \text{for } \rho = 1 - p \in (0, 1].$$

## Acknowledgements

I especially thank Barbara Niethammer for her hospitality and many discussions during my visit to Berlin. This work was supported by the Deutsche Forschungsgemeinschaft in the form of a Mercator professorship at Humboldt University, by the Max Planck Institute for Mathematics in the Sciences in Leipzig, and by the Institute for Mathematical Sciences at the National University of Singapore. This material is based upon work supported by the National Science Foundation under grant DMS 03-05985.

## References

1. N. D. ALIKAKOS, P. W. BATES, AND X. CHEN, *Convergence of the Cahn-Hilliard equation to the Hele-Shaw model*, Arch. Rational Mech. Anal., 128 (1994), pp. 165–205.
2. N. D. ALIKAKOS, G. FUSCO, AND G. KARALI, *Ostwald ripening in two dimensions—the rigorous derivation of the equations from the Mullins-Sekerka dynamics*, J. Differential Equations, 205 (2004), pp. 1–49.
3. S. ALLEN AND J. CAHN, *Microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta Metallurgica, 27 (1979), pp. 1085–1095.
4. G. ANDREWS AND J. M. BALL, *Asymptotic behaviour and changes of phase in one-dimensional nonlinear viscoelasticity*, J. Differential Equations, 44 (1982), pp. 306–341. Special issue dedicated to J. P. LaSalle.
5. G. S. BALES AND A. ZANGWILL, *Morphological instability of a terrace edge during step-flow growth*, Phys. Rev. B, 41 (1990), pp. 5500–5508.
6. N. H. BINGHAM, C. M. GOLDIE, AND J. L. TEUGELS, *Regular variation*, vol. 27 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1987.
7. W. K. BURTON, N. CABRERA, AND F. C. FRANK, *The growth of crystals and the equilibrium structure of their surfaces*, Philos. Trans. Roy. Soc. London. Ser. A., 243 (1951), pp. 299–358.
8. G. CAGINALP, *Stefan and Hele-Shaw type models as asymptotic limits of the phase-field equations*, Phys. Rev. A (3), 39 (1989), pp. 5887–5896.
9. J. CARR AND R. PEGO, *Invariant manifolds for metastable patterns in  $u_t = \epsilon^2 u_{xx} - f(u)$* , Proc. Roy. Soc. Edinburgh Sect. A, 116 (1990), pp. 133–160.
10. ———, *Self-similarity in a coarsening model in one dimension*, Proc. Roy. Soc. London Ser. A, 436 (1992), pp. 569–583.
11. J. CARR AND R. L. PEGO, *Metastable patterns in solutions of  $u_t = \epsilon^2 u_{xx} - f(u)$* , Comm. Pure Appl. Math., 42 (1989), pp. 523–576.
12. J. CARR AND O. PENROSE, *Asymptotic behaviour of solutions to a simplified Lifshitz-Slyozov equation*, Phys. D, 124 (1998), pp. 166–176.
13. X. CHEN, *Generation, propagation, and annihilation of metastable patterns*, J. Differential Equations, 206 (2004), pp. 399–437.

14. S. DAI, *Universal bounds on coarsening rates for some models of phase transitions*, PhD thesis, University of Maryland, College Park, 2005.
15. S. DAI AND R. L. PEGO, *Universal bounds on coarsening rates for mean-field models of phase transitions*, *SIAM J. Math. Anal.*, 37 (2005), pp. 347–371 (electronic).
16. J. DAVIS, *The new diamond age*, *Wired*, 11.09 (2003).  
<http://www.wired.com/wired/archive/11.09/diamond.html>.
17. B. DERRIDA, *Coarsening phenomena in one dimension*, in *Complex systems and binary networks* (Guanajuato, 1995), vol. 461 of *Lecture Notes in Phys.*, Springer, Berlin, 1995, pp. 164–182.
18. J.-P. ECKMANN AND J. ROUGEMONT, *Coarsening by Ginzburg-Landau dynamics*, *Comm. Math. Phys.*, 199 (1998), pp. 441–470.
19. W. FELLER, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971.
20. P. C. FIFE, *Dynamics of internal layers and diffusive interfaces*, vol. 53 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1988.
21. G. FUSCO AND J. K. HALE, *Slow-motion manifolds, dormant instability, and singular perturbations*, *J. Dynam. Differential Equations*, 1 (1989), pp. 75–94.
22. T. GALLAY AND A. MIELKE, *Convergence results for a coarsening model using global linearization*, *J. Nonlinear Sci.*, 13 (2003), pp. 311–346.
23. B. GIRON, B. MEERSON, AND P. V. SASOROV, *Weak selection and stability of localized distributions in ostwald ripening*, *Phys. Rev. E*, 58 (1998), pp. 4213–4216.
24. K. GLASNER, *A diffuse interface approach to Hele-Shaw flow*, *Nonlinearity*, 16 (2003), pp. 49–66.
25. J. K. HALE AND P. MASSAT, *Asymptotic behavior of gradient-like systems*, in *Dynamical systems II* (Proc. 2nd Intl. Symp., Gainesville, Fla., 1981), A. R. Bednarek and L. Cesari, eds., New York, 1982, Academic Press Inc., pp. 85–101.
26. H.-C. JEONG AND E. WILLIAMS, *Steps on surfaces: experiment and theory*, *Surf. Sci. Reports*, 34 (1999), pp. 175–294.
27. R. V. KOHN AND F. OTTO, *Upper bounds on coarsening rates*, *Comm. Math. Phys.*, 229 (2002), pp. 375–395.
28. R. V. KOHN AND X. YAN, *Upper bound on the coarsening rate for an epitaxial growth model*, *Comm. Pure Appl. Math.*, 56 (2003), pp. 1549–1564.
29. R. V. KOHN AND X. YAN, *Coarsening rates for models of multicomponent phase separation*, *Interfaces Free Bound.*, 6 (2004), pp. 135–149.
30. B. LI AND J.-G. LIU, *Epitaxial growth without slope selection: energetics, coarsening, and dynamic scaling*, *J. Nonlinear Sci.*, 14 (2004), pp. 429–451 (2005).
31. I. M. LIFSHITZ AND V. V. SLYOZOV, *The kinetics of precipitation from supersaturated solid solutions*, *J. Phys. Chem. Solids*, 19 (1961), pp. 35–50.
32. R. MENDOZA, J. ALKEMPER, AND P. W. VOORHEES, *The morphological evolution of dendritic microstructures during coarsening*, *Metall. and Mater. Trans.*, 34A (2003), pp. 481–489.

33. G. MENON AND R. L. PEGO, *The scaling attractor and ultimate dynamics for Smoluchowski's coagulation equation*, J. Nonlinear Sci., to appear.
34. G. MENON AND R. L. PEGO, *Approach to self-similarity in Smoluchowski's coagulation equations*, Comm. Pure Appl. Math., 57 (2004), pp. 1197–1232.
35. D. MOLDOVAN AND L. GOLUBOVIC, *Interfacial coarsening dynamics in epitaxial growth with slope selection*, Phys. Rev. E, 61 (2000), pp. 6190–6214.
36. T. NAGAI AND K. KAWASAKI, *Statistical dynamics of interacting kinks II*, Physica A, 134 (1986), pp. 483–521.
37. B. NIETHAMMER, *Derivation of the LSW-theory for Ostwald ripening by homogenization methods*, Arch. Ration. Mech. Anal., 147 (1999), pp. 119–178.
38. ———, *The LSW model for Ostwald ripening with kinetic undercooling*, Proc. Roy. Soc. Edinburgh Sect. A, 130 (2000), pp. 1337–1361.
39. B. NIETHAMMER AND F. OTTO, *Ostwald ripening: the screening length revisited*, Calc. Var. Partial Differential Equations, 13 (2001), pp. 33–68.
40. B. NIETHAMMER AND R. L. PEGO, *Non-self-similar behavior in the LSW theory of Ostwald ripening*, J. Statist. Phys., 95 (1999), pp. 867–902.
41. ———, *On the initial-value problem in the Lifshitz-Slyozov-Wagner theory of Ostwald ripening*, SIAM J. Math. Anal., 31 (2000), pp. 467–485 (electronic).
42. ———, *Well-posedness for measure transport in a family of nonlocal domain coarsening models*, Indiana Univ. Math. J., 54 (2005), pp. 499–530.
43. A. NOVICK-COHEN AND R. L. PEGO, *Stable patterns in a viscous diffusion equation*, Trans. Amer. Math. Soc., 324 (1991), pp. 331–351.
44. F. OTTO, *Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory*, Arch. Rational Mech. Anal., 14 (1998), pp. 63–103.
45. F. OTTO, P. PENZLER, A. RÄTZ, T. RUMP, AND A. VOIGT, *A diffuse-interface approximation for step flow in epitaxial growth*, Nonlinearity, 17 (2004), pp. 477–491.
46. R. L. PEGO, *Front migration in the nonlinear Cahn-Hilliard equation*, Proc. Roy. Soc. London Ser. A, 422 (1989), pp. 261–278.
47. R. L. PEGO, *Stabilization in a gradient system with a conservation law*, Proc. Amer. Math. Soc., 114 (1992), pp. 1017–1024.
48. J. RUBINSTEIN, P. STERNBERG, AND J. B. KELLER, *Fast reaction, slow diffusion, and curve shortening*, SIAM J. Appl. Math., 49 (1989), pp. 116–133.
49. E. SENETA, *Regularly varying functions*, Springer-Verlag, Berlin, 1976. Lecture Notes in Mathematics, Vol. 508.
50. C. WAGNER, *Theorie der alteration von niederschlägen durch umlösen*, Z. Elektrochemie, 65 (1961), pp. 581–594.
51. S. J. WATSON, *Coarsening dynamics of growing faceted crystal surfaces: the annealing to growth transition*. submitted.