Contents

Preface

Statistical physics is an unusual branch of physics because it is not really a well-defined field in a formal sense, but rather, statistical physics is a viewpoint — indeed, the most appropriate viewpoint — to investigate systems with many degrees of freedom. Part of the appeal of statistical physics is that it can be applied to a disparate range of systems that are in the mainstream of physics, as well as to problems that might appear to be outside physics, such as econophysics, quantitative biology, and social organization phenomena. Many of the basic features of these systems involve fluctuations or explicit time evolution rather than equilibrium distributions. Thus the approaches of non-equilibrium statistical physics are needed to discuss such systems.

While the tools of equilibrium statistical physics are well-developed, the statistical description of systems that are out of equilibrium is still relatively primitive. In spite of more than a century of effort in constructing an over-arching approach for non-equilibrium phenomena, there still does not exist canonical formulations, such as the Boltzmann factor or the partition function in equilibrium statistical physics. At the present time, some of the most important theoretical approaches for non-equilibrium systems are either technical, such as deriving hydrodynamics from the Boltzmann equation, or somewhat removed from the underlying phenomena that are being described, such as non-equilibrium thermodynamics.

Because of this disconnect between fundamental theory and applications, our view is that it is more instructive to illustrate non-equilibrium statistical physics by presenting a number of current and paradigmatic examples of systems that are out of equilibrium, and to elucidate, as completely as possible, the range of techniques available to solve these systems. By this approach, we believe that readers can gain general insights more quickly compared to formal approaches, and, further, will be well-equipped to understand many other topics in non-equilibrium statistical physics. We have attempted to make our treatment as self-contained and user-friendly as possible, so that an interested reader can work through the book without encountering unresolved methodological mysteries or hidden calculational pitfalls. Thus while much of the material is mathematical in nature, we have tried to present it as pedagogically as possible. Our target audience is graduate students with a one-course background in equilibrium statistical physics. Each of the main chapters is intended to be self-contained. We also made an effort to supplement the chapters with research exercises and open questions, in the hopes of stimulating further research.

The specific examples presented in this book are primarily based on irreversible stochastic processes. This branch of statistical physics is in many ways the natural progression of kinetic theory that was initially used to describe dynamics of simple gases and fluids. We will discuss the development of basic kinetic approaches to more complex and contemporary systems. Among the large menu of stochastic and irreversible processes, we chose the ones that we consider to be among the most important and most instructive in leading to generic understanding. Our main emphasis is on exact analytical results, but we also send time developing heuristic and scaling methods. We largely avoid presenting numerical simulation results because these are less definitive and instructive than analytical results. An appealing (at least to us) aspect of these examples is that they are broadly accessible. One needs little background to appreciate the systems being studied and the ideas underlying the methods of solution. Many of these systems naturally suggest new and non-trivial questions that an interested reader can easily pursue.

We begin our exposition with a few "aperitifs" — an abbreviated qualitative discussion of basic problems and a general hint at the approaches that are available to solve these systems. Chapter 2 provides a basic introduction to diffusion phenomena because of the central role played by diffusion in many non-equilibrium statistical systems. These preliminary chapters serve as an introduction to the rest of the book.

The main body of the book is then devoted to working out specific examples. In the next three chapters, we discuss the fundamental kinetic processes of aggregation, fragmentation, and adsorption (chapters 3–5).

Aggregation is the process by which two clusters irreversibly combine in a mass-conserving manner to form a larger cluster. This classic process that very nicely demonstrates the role of conservation laws, the utility of exact solutions, the emergence of scaling in cluster-size distributions, and the power of heuristic derivations. Many of these technical lessons will be applied throughout this book.

We then turn to the complementary process of fragmentation, which involves the repeated breakup of an element into smaller fragments. While this phenomenon again illustrates the utility of exact and scaling solutions, fragmentation also exposes important new concepts such as multiscaling, lack of self-averaging, and methods such as traveling waves and velocity selection. All of these are concepts that are used extensively in non-equilibrium statistical physics. We then discuss the phenomenon of irreversible adsorption where, again, the exact solutions of underlying master equations for the occupancy probability distributions provides a comprehensive picture of the basic phenomena.

The next two chapters (6 & 7) discuss the time evolution of systems that involve the competition between multiple phases. We first treat classical spin systems, in particular, the kinetic Ising model and the voter model. The kinetic Ising model occupies a central role in statistical physics because of its broad applicability to spins systems and many other dynamic critical phenomena. The voter model is perhaps not as well known in the physics literature, but it is an even simpler model of an evolving spin system that is exactly soluble in all dimensions. In chapter 7, we study phase ordering kinetics. Here the natural descriptions are in terms of continuum differential equations, rather than master equations.

In chapter 8, we discuss collision-driven phenomena. Our aim is to present the Boltzmann equation in the context of explicitly soluble examples. These include traffic models, and aggregation and annihilation processes in which the particles move at constant velocity between collisions. The final chapter (#9) presents several applications to contemporary problems, such as the structure of growing networks and models of self-organized criticality.

In an appendix, we present the fundamental techniques that are used throughout our book. These include various types of integral transforms, generating functions, asymptotic analysis, extreme statistics, and scaling approaches. Each of these methods is explain fully upon its first appearance in the book itself, and the appendix is a brief compendium of these methods that can be used either as a reference or a study guide, depending on one's technical preparation.

Chapter 1

APERITIFS

Broadly speaking, non-equilibrium statistical physics describes the time-dependent evolution of many-particle systems. The individual particles are elemental interacting entities which, in some situations, can change in the process of interaction. In the most interesting cases, interactions between particles are strong and hence the deterministic description of even few-particle systems are beyond the reach of exact theoretical approaches. On the other hand, many-particle systems often admit an analytical statistical description when their number becomes large and in that sense they are **simpler** than few-particle systems. This feature has several different names—the law of large numbers, ergodicity, *etc.*—and it is one of the reasons for the spectacular successes of statistical physics and probability theory.

Non-equilibrium statistical physics is quite different from other branches of physics, such as the 'fundamental' fields of electrodynamics, gravity, and elementary-particle physics that involve a reductionist description of few-particle systems, and applied fields, such as hydrodynamics and elasticity that are primarily concerned with the consequences of fundamental governing equations. Some of the key and distinguishing features of non-equilibrium statistical physics include:

- no basic equations (like Maxwell equations in electrodynamics or Navier-Stokes equations in hydrodynamics) from which the rest follows;
- intermediate between fundamental applied physics;
- the existence of common underlying techniques and concepts in spite of the wide diversity of the field;
- non-equilibrium statistical physics naturally leads to the creation of methods that are quite useful in applications far removed from physics (for example the Monte Carlo method and simulated annealing).

Our guiding philosophy is that in the absence of underlying principles or governing equations, non-equilibrium statistical physics should be oriented toward explicit and illustrative examples rather than attempting to develop a theoretical formalism that is still incomplete.

Let's start by looking briefly at the random walk to illustrate a few key ideas and to introduce several useful analysis tools that can be applied to more general problems.

1.1 Diffusion

For the symmetric diffusion on a line, the probability density

$$\text{Prob}\,[\text{particle} \in (x, x+dx)] \equiv P(x,t)\,dx \tag{1.1}$$

satisfies the diffusion equation

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2}. \tag{1.2}$$

As we discuss soon, this equation describes the continuum limit of an unbiased random walk. The diffusion equation must be supplemented by an initial condition that we take to be $P(x,0) = \delta(x)$, corresponding to a walk that starts at the origin.

Dimensional Analysis

Let's pretend that we don't know how to solve (1.2) and try to understand the behavior of the walker without explicit solution. What is the mean displacement? There is no bias, so clearly

$$\langle x\rangle \equiv \int_{-\infty}^{\infty} x\, P(x,t)\, dx = 0\,.$$

The next moment, the mean square displacement,

$$\langle x^2\rangle \equiv \int_{-\infty}^{\infty} x^2\, P(x,t)\, dx$$

is non-trivial. Obviously, it should depend on the diffusion coefficient D and time t. We now apply dimensional analysis to determine these dependences. If L denotes the unit of length and T denotes the time unit, then from (1.2) the dimensions of $\langle x^2\rangle, D$, and t are

$$[\langle x^2\rangle] = L^2\,, \quad [D] = L^2/T\,, \quad [t] = T\,.$$

The ratio $\langle x^2\rangle/Dt$ is dimensionless and thus be constant, as a dimensionless quantity cannot depend on dimensional quantities. Hence

$$\langle x^2\rangle = C\times Dt. \tag{1.3}$$

Equation (1.3) is one of the central results in non-equilibrium statistical physics, and we derived it using just dimensional analysis! To determine the numerical constant $C = 2$ in (1.3) one must work a bit harder (*e.g.*, by solving (1.2), or by multiplying Eq. (1.2) by x^2 and integrating over the spatial coordinate to give $\frac{d}{dt}\,\langle x^2\rangle = 2D$). We shall therefore use the power of dimensional analysis whenever possible.

Scaling

Let's now apply dimensional analysis to the probability density $P(x,t|\,D)$; here D is explicitly displayed to remind us that the density does depend on the diffusion coefficient. Since $[P] = L^{-1}$, the quantity $\sqrt{Dt}\,P(x,t|\,D)$ is dimensionless, so it must depend on dimensionless quantities only. From variables x, t, D we can form a single dimensionless quantity $x/\sqrt{Dt}$. Therefore the most general dependence of the density on the basic variables that is allowed by dimensional analysis is

$$P(x,t) = \frac{1}{\sqrt{Dt}}\,\mathcal{P}(\xi)\,, \qquad \xi = \frac{x}{\sqrt{Dt}}\,. \tag{1.4}$$

The density depends on a single *scaling variable* rather than on two basic variables x and t. This remarkable feature greatly simplifies analysis of the typical partial differential equations that describe non-equilibrium systems. Equation (1.4) is often referred to as the *scaling ansatz*. Finding the right scaling ansatz for a physical problem often represents a large step step toward a solution. For the diffusion equation (1.2), substituting in the ansatz (1.4) reduces this partial differential equation to the ordinary differential equation

$$2\mathcal{P}'' + \xi\mathcal{P}' + \mathcal{P} = 0\,.$$

Integrating twice and invoking both symmetry ($\mathcal{P}'(0) = 0$) and normalization, we obtain $\mathcal{P} = (4\pi)^{-1/2}\, e^{-\xi^2/4}$, and finally the Gaussian probability distribution

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}}\,\exp\left\{-\frac{x^2}{4Dt}\right\}\,. \tag{1.5}$$

In this example, the scaling form was rigorously derived from simple dimensional reasoning. In more complicated situations, arguments in favor of scaling are less rigorous, and scaling is usually achieved only in some asymptotic limit . The above example where scaling applies for all t is an exception; for the diffusion equation with an initial condition on a finite rather than a point support, scaling holds only in the limit $x, t\to\infty$ with the scaling variable ξ kept finite. Nevertheless, we shall see that where applicable, scaling provides a significant step toward the understanding of a problem.

Renormalization

The strategy of the renormalization group method is to understand the behavior on large 'scale'—here large time—iteratively in terms of the behavior on smaller scales. For the diffusion equation, we start with identity

$$P(x,2t) = \int_{-\infty}^{\infty} P(y,t)\, P(x-y,t)\, dy \tag{1.6}$$

that reflects the fact that the random walk is a Markov process. Namely, to reach x at time $2t$, the walk first reaches some intermediate point y at time t and then completes the journey to y in the remaining time t. (Equation (1.6) is also the basis for the path integral treatment of diffusion processes but we will not delve into this subject here.)

The convolution form of Eq. (1.6) calls out for applying the Fourier transform,

$$\hat{P}(k,t) = \int_{-\infty}^{\infty} e^{ikx}\, P(x,t)\, dx\,, \tag{1.7}$$

that recasts (1.6) into the algebraic relation $\hat{P}(k,2t) = [\hat{P}(k,t)]^2$. The scaling form (1.4) shows that $\hat{P}(k,t) = \hat{\mathcal{P}}(\kappa)$ with $\kappa = k\sqrt{Dt}$, so the renormalization group equation is

$$\hat{\mathcal{P}}(\sqrt{2}\,\kappa) = [\hat{\mathcal{P}}(\kappa)]^2\,.$$

Taking logarithms and defining $z \equiv \kappa^2, Q(z) \equiv \ln \hat{\mathcal{P}}(\kappa)$, we arrive at $Q(2z) = 2Q(z)$, whose solution is $Q(z) = -Cz$, or $\hat{P}(k,t) = e^{-2k^2 Dt}$. (The constant $C = 2$ may be found, *e.g.*, by expanding (1.7) for small k, $\hat{P}(k,t) = 1 - k^2\langle x^2\rangle$, and recalling that $\langle x^2\rangle = 2Dt$). Performing the inverse Fourier transform we recover (1.5). Thus the Gaussian probability distribution represents an exact solution to a renormalization group equation. Our derivation shows that the renormalization group is ultimately related to scaling.

1.2 Single-Species Annihilation/Coalescence

In non-equilibrium statistical physics, we study systems that contain a macroscopic number of interacting particles. To understand collective behaviors it is useful to ignore complications resulting from finiteness, *i.e.*, to focus on situations when the number of particles is infinite. Perhaps the simplest interacting infinite-particle systems of this kind are *single-species annihilation*, where particles diffuse freely and annihilate instantaneously upon contact, and *single-species coalescence*, where the reactants merge upon contact. These processes have played an important role in development of non-equilibrium statistical physics and they provide excellent illustrations of techniques that can be applied to other infinite-particle systems.

The annihilation process is symbolically represented by the reaction scheme

$$A + A \longrightarrow \emptyset\,, \tag{1.8}$$

while the coalescence reaction is represented by

$$A + A \longrightarrow A. \tag{1.9}$$

The density $n(t)$ of A particles for both reactions obviously decays with time; the question is: how?

Hydrodynamics

In the hydrodynamic approach, one assumes that the reactants are perfectly mixed at all times. This means that the density at every site is the same and that every particle has the same probability to react at the next instant. In this well-mixed limit, and also assuming the continuum limit, the global particle density n for both annihilation and coalescence decays with time according to the *rate equation*

$$\frac{dn}{dt} = -Kn^2\,. \tag{1.10}$$

This equation reflects that fact that two particles are needed for a reaction to occur and the probability for two particle to be at the same location is proportional to the density squared. Here K is the reaction rate that describes the propensity for two diffusing particles to interact; the computation of this rate requires a detailed microscopic treatment (see chapter 4). The rate equation (1.10) is a typical hydrodynamic-like equation whose solution is

$$n(t) = \frac{n_0}{1 + Kn_0 t} \sim (Kt)^{-1}\,. \tag{1.11}$$

However, simulations show more interesting long-time behaviors that depends on the spatial dimension d:

$$n(t) \sim \begin{cases} t^{-1/2} & d = 1; \\ t^{-1}\,\ln t & d = 2; \\ t^{-1} & d > 2. \end{cases} \tag{1.12}$$

The sudden change at $d_c = 2$ illustrates the important notion of the critical dimension: above d_c, the rate equation leads to asymptotically correct behavior; below d_c, the rate equation is wrong; at d_c, the rate equation approach is almost correct—it typically is in error by a logarithmic correction term.

To obtain a complete theory of the reaction, one might try to write formally exact equations for correlation functions. That is, if $\rho(\mathbf{r}, t)$ is the microscopic density, the true dynamical equation for $n(t) \equiv \langle\rho(\mathbf{r}, t)\rangle$ involves the second-order correlators $\langle\rho(\mathbf{r}, t)\rho(\mathbf{r}', t)\rangle$. Then an equation for the second-order correlation functions involves third-order correlators, *etc.* These equations are hierarchical and the only way to proceed is to impose some sort of closure scheme in which higher-order correlators are factorized in terms of lower-order correlators. In particular, the hydrodynamic equation (1.10) is recovered if we assume that second-order correlators factorize; that is, $\langle\rho(\mathbf{r}, t)\rho(\mathbf{r}', t)\rangle = \langle\rho(\mathbf{r}, t)\rangle\langle\rho(\mathbf{r}', t)\rangle = n(t)^2$. Thus Eq. (1.10) is the factorized version of the Boltzmann equation for the annihilation process (1.8). Attempts to describe this reaction scheme more faithfully by higher-order correlators have not been fruitful. Thus the revered kinetic theory approach is helpless for the innocent-looking process (1.8)! Let's try some other approaches.

Dimensional Analysis

Let's determine the dependence of the rate K on fundamental parameters of the reaction, *i.e.*, on the diffusion coefficient D of the reactants and radius R of each particle. From Eq. (1.10), $[K] = L^d/T$, and the only possible dependence is [1]

$$K = DR^{d-2} \tag{1.13}$$

Using (1.13) in (1.10) and solving this equation yields

$$n(t) \sim \frac{1}{R^{d-2}Dt}\,. \tag{1.14}$$

We anticipate that the density ought to decay more quickly when the radius of the particles is increased. According to (1.14), this is true only when $d > 2$. Thus the rate equation could be correct in this regime. Surprisingly, however, the reaction rate is not proportional to the cross-sectional area, R^{d-1}, but rather to R^{d-2}; this feature stems from the vagaries of diffusive motion. For $d = 2$, the decay is independent of the size of particles—already a bit of a surprising result. However, for $d < 2$, we obtain the obviously wrong result that the density decays more slowly if particles are larger.

The density is actually *independent* of R for $d < 2$. This fact is easy to see for $d = 1$ because all that matters is the spacing between particles. If we now seek, on dimensional grounds, the density in the R-independent form $n(D, t)$, we find that the only possibility is $n \propto (Dt)^{-d/2}$ in agreement with prediction of (1.12) in one dimension. In the context of the reaction rate, this slow decay is equivalent to a reaction rate that decreases with time. We will return to this point in the next chapter.

[1] Here we omit a numerical factor of order one; in the future, we shall often ignore such factors without explicit warning.

Heuristic Arguments

Dimensional analysis often gives correct dependences but does not really explain why these behaviors are correct. For the annihilation process (1.8), we can understand the one-dimensional asymptotic, $n \sim (Dt)^{-1/2}$, in a physical way by using a basic feature (1.3) of random walks: in a time interval $(0, t)$, each particle explores the region $\ell \sim \sqrt{Dt}$, and therefore a typical separation between surviving particles is of order ℓ, from which $n \sim \ell^{-1} \sim (Dt)^{-1/2}$ follows.

Guided by this understanding, let's try to understand (1.12) for all dimensions. First, we slightly modify the process so that particles undergo diffusion on a lattice in d dimensions (lattice spacing plays the role of the radius). What is the average number of sites $\mathcal{N}$ visited by a random walker after N steps? This question has a well-known and beautiful answer:

$$\mathcal{N} \sim \begin{cases} N^{1/2} & d = 1; \\ N/\ln N & d = 2; \\ N & d > 2. \end{cases} \tag{1.15}$$

With a little contemplation, one should be convinced the density in single-species annihilation scales as the inverse of the average number of sites visited by a random walker; if there is more that one particle in the visited region, it should have been annihilated previously. Thus (1.15) is essentially equivalent to (1.12).

Exact Solution in One Dimension

The diffusion-controlled annihilation process admits an exact solution in one dimension. This is an exceptional feature—most infinite-particle systems cannot be solved even in one dimension. Moreover, for these solvable cases, we can usually compute only a limited number of quantities. For one-dimensional annihilation, for example, while the density is known exactly, the distribution of distances ℓ between adjacent particles $P(\ell, t)$ is unknown even in the scaling limit $\ell \to \infty$ and $t \to \infty$, with $\xi = \ell/\sqrt{Dt}$ being finite. Although numerical simulations strongly indicate that the interval length distribution approaches the scaling form, $P(\ell, t) \to (Dt)^{-1/2}\mathcal{P}(\xi)$, nobody yet knows how to compute the scaled length distribution $\mathcal{P}(\xi)$.

Exact results for the diffusion-controlled annihilation process will be presented later when we develop the necessary technical tools. However to illustrate a simple exact solution, let's consider diffusion-controlled coalescence, $A + A \longrightarrow A$, that *is* readily soluble in one dimension because it can be reduced to a two-particle problem. To compute the density it is convenient to define particle labels so that in each collision the left particle disappears and the right particle survives. Then to compute the survival probability of a test particle we may ignore all particles to the left. Such a reduction of the original two-sided problem to a one-sided one is extremely helpful. Furthermore, only the closest particle to the right of the test particles is relevant—the right neighbor can merge with other particles further to the right; however, these reactions never affect the fate of the test particle. Thus the system reduces to a soluble two-particle problem.

The interparticle distance between the test particle and its right neighbor undergoes diffusion with diffusivity $2D$ because the spacing diffuses at twice the rate of each particle. Consequently, the probability density $\rho(\ell, t)$ that the test particle is separated by distance ℓ from its right neighbor satisfies the diffusion equation subject to the absorbing boundary condition:

$$\frac{\partial \rho}{\partial t} = 2D\frac{\partial^2 \rho}{\partial \ell^2}\,, \qquad \rho(0, t) = 0. \tag{1.16}$$

The solution (1.16) for an arbitrary initial condition $\rho_0(\ell)$ is

$$\begin{aligned} \rho(\ell, t) &= \frac{1}{\sqrt{8\pi Dt}} \int_0^\infty \rho_0(y)\, \left[e^{-(\ell-y)^2/8Dt} - e^{-(\ell+y)^2/8Dt}\right] dy \\ &= \frac{1}{\sqrt{2\pi Dt}}\, \exp\left(-\frac{\ell^2}{8Dt}\right) \int_0^\infty \rho_0(y)\, \exp\left(-\frac{y^2}{8Dt}\right) \sinh\left(\frac{\ell y}{4Dt}\right) dy\,. \end{aligned} \tag{1.17}$$

In the first line, the solution is expressed as the superposition of a Gaussian and an image anti-Gaussian that automatically satisfies the absorbing boundary condition. In the long time limit, the integral on the

second line tends to $\frac{\ell}{4Dt}\int_0^\infty dy\,\rho_0(y)\,y = \frac{\ell}{4Dtn_0}$. Therefore

$$\rho(\ell,t) \to \frac{\ell}{4Dtn_0\sqrt{2\pi Dt}}\,\exp\left(-\frac{\ell^2}{8Dt}\right),$$

so that the survival probability is

$$S(t) = \int_0^\infty \rho(\ell,t)\,d\ell \to n_0^{-1}(2\pi Dt)^{-1/2}\,,$$

and the density $n(t) = n_0\,S(t)$ decays as

$$n(t) \to (2\pi Dt)^{-1/2} \quad \text{when} \quad t\to\infty\,. \tag{1.18}$$

To summarize, the interval length distribution $P(\ell,t)$ is just the probability density $\rho(\ell,t)$ conditioned on the survival of the test particle. Hence

$$P(\ell,t) \equiv \frac{\rho(\ell,t)}{S(t)} \to \frac{\ell}{4Dt}\,\exp\left(-\frac{\ell^2}{8Dt}\right).$$

The average interparticle spacing grows as $\sqrt{Dt}$ which equivalent to the particle density decaying as $1/\sqrt{Dt}$.

1.3 Two-Species Annihilation

Consider two diffusing species A and B which are initially distributed at random with equal concentrations: $n_A(0) = n_B(0) = n_0$. When two particles of opposite species approach within the reaction radius, they immediately annihilate:

$$A + B \longrightarrow \emptyset\,. \tag{1.19}$$

For this reaction, the density decreases as

$$n(t) \sim \begin{cases} t^{-d/4} & d \le 4; \\ t^{-1} & d > 4, \end{cases} \tag{1.20}$$

as $t\to\infty$, so the critical dimension is $d_c = 4$. This result shows that hydrodynamic description is wrong even in the most relevant three-dimensional case.

In this striking example, neither a hydrodynamic description (that gives $n\sim t^{-1}$) nor dimensional analysis can explain the decay of the density. Here, a simple heuristic argument helps us determine the density decay of Eq. (1.20). To understand why the naive approaches fail, consider a snapshot of a two-dimensional system at some time $t \gg 1$ (Fig. fig-aper-snapshot). We see that the system spontaneously organizes into a mosaic of alternating domains. Because of this organization, annihilation can occur only along domain boundaries rather than throughout the system. This screening effect explains why the density is much larger than in the hydrodynamic picture where particles are assumed to be well-mixed.

To turn this picture into a semi-quantitative estimate for the density, note that in a spatial region of linear size ℓ, the initial number of A particles is $N_A = n_0\ell^d \pm (n_0\ell)^{d/2}$ and similarly for B particles. Here the $\pm$ term signifies that the particle number in a finite region is a stochastic variable that typically fluctuates in a range of order $(n_0\ell)^{d/2}$ about the mean value $n_0\ell^d$. The typical value of the difference $N_A - N_B$ for this d-dimensional region

$$N_A - N_B = \pm(n_0\ell)^{d/2},$$

arises because of initial fluctuations and is not affected by annihilation events. Therefore after the minority species in a given region is eliminated, the local density becomes $n\sim (n_0\ell)^{d/2}/\ell^d$. Because of the diffusive spreading (1.3), the average domain size scales as $\ell\sim\sqrt{Dt}$, and thus $n\sim\sqrt{n_0}\,(Dt)^{-d/4}$. Finally, notice that the density decay cannot be obtained by dimensional analysis alone because now are there at least two independent length scales, the domain size $\sqrt{Dt}$ and the interparticle spacing. Additional physical input, here in the form of the domain picture, is needed to obtain $n(t)$.

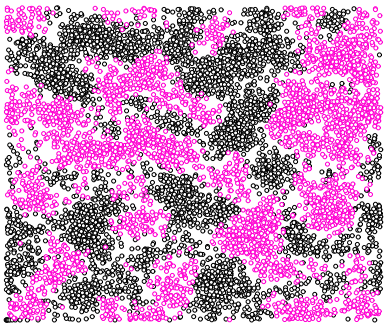


Figure 1.1: Snapshot of the particle positions in two-species annihilation in two dimensions.

Notes

There is a large literature on the topics discussed in this introductory chapter. Among the great many books on random walks we mention 3; 8 which contain numerous further references. Dimensional analysis and scaling are especially popular in hydrodynamics, see e.g. excellent reviews in 4 and the classical book by Barenblatt 1 which additionally emphasizes the connection of scaling, especially intermediate asymptotics, and the renormalization group. This latter connection has been further explored by many authors, particularly Goldenfeld and co-workers (see 6). Kinetics of single-species and two-species annihilation processes were understood in a pioneering works of Zeldovich, Ovchinnikov, Burlatskii, Toussaint, Wilczek, Bramson, Lebowitz, and many others; a review of this work is given in 5; 2. One-dimensional diffusion-controlled coalescence process is one of the very few examples which can justifiably be called "completely solvable"; numerous exact solutions of this model (and generalizations thereof) found by ben-Avraham, Doering, and others are presented in Ref. 2.

Chapter 2

RANDOM WALK/DIFFUSION

Because the random walk and its continuum diffusion limit underlie so many fundamental processes in non-equilibrium statistical physics, we give a brief introduction to this central topic. There are several complementary ways to describe random walks and diffusion, each with their own advantages.

2.1 Langevin Equation

We begin with the phenomenological Langevin equation that represents a minimalist description for the stochastic motion of a random walk. We mostly restrict ourselves to one dimension, but the generalization to higher dimensions is straightforward. Random walk motion arises, for example, when a microscopic bacterium is placed in a fluid. The bacterium is constantly buffeted on a very short time scale by the random collisions with fluid molecules. In the Langevin approach the effect of these rapid collisions is represented by an effective, but stochastic, external force $\eta(t)$. On the other hand, if the bacterium had a non-zero velocity in the fluid, there would be a systematic frictional force proportional to the velocity that would bring the bacterium to rest. Under the influence of these two forces, Newton's second law for the motion of the bacterium leads to the Langevin equation

$$m\frac{dv}{dt} = -\gamma v + \eta(t). \tag{2.1}$$

This equation is very different from the deterministic equation of motion that one normally encounters in mechanics. Because the stochastic force is so rapidly changing with time, the actual trajectory of the particle contains too much information. The velocity changes every time there is a collision between the bacterium and a fluid molecule; for a particle of linear dimension 1μm, there are of the order of 10^{20} collisions per second and it is pointless to follow the motion on such a short time scale. For this reason, it is more meaningful physically to study the trajectory that is averaged over longer times. To this end, we need to specify the statistical properties of the random force. Because the force is a result of molecular collisions, it is natural to assume that the force $\eta(t)$ is a random function of time with zero mean, $\langle \eta(t) \rangle = 0$. Here the angle brackets denote the time average. Because of the rapidly fluctuating nature of the force, we also assume that there is no correlation between the force at two different times, so that $\langle \eta(t)\eta(t') \rangle = 2D\gamma^2\delta(t-t')$. As a result, the product of the forces at two different times has a mean value of zero. However, the *mean-square* force at any time has the value D. This statement merely states that the average *magnitude* of the force is well-defined.

In the limit where the mass of the bacterium is sufficiently small that it may be neglected, we obtain an even simpler equation for the position of the bacterium:

$$\frac{dx}{dt} = \frac{1}{\gamma}\eta(t) \equiv \xi(t). \tag{2.2}$$

In this limit of no inertia ($m = 0$) the instantaneous velocity equals the force. In spite of this strange feature, Eq. (2.2) has a simple interpretation—the change in position is a randomly fluctuating variable. This corresponds to a naive view of what a random walk actually does; at each step the position changes by a random amount.

One of the advantages of the Langevin equation description is that average values of the moments of the position can be obtained quite simply. Thus formally integrating Eq. (2.1), we obtain

$$x(t) = \int_0^t \xi(t')\, dt'. \tag{2.3}$$

Because $\langle \xi(t) \rangle = 0$, then $\langle x(t) \rangle = 0$. However, the mean-square displacement is non-trivial. Formally,

$$\langle x(t)^2 \rangle = \int_0^t \int_0^t \langle \xi(t')\xi(t'') \rangle \, dt'\, dt''. \tag{2.4}$$

Using $\langle \xi(t)\xi(t') \rangle = 2D\delta(t-t')$, it immediately follows that $\langle x(t)^2 \rangle = 2Dt$. Thus we recover the classical result that the mean-square displacement grows linearly in time. Furthermore, we can identify D as the diffusion coefficient. The dependence of the mean-square displacement can also be obtained by dimensional analysis of the Langevin equation. Because the delta function $\delta(t)$ has units of $1/t$ (since the integral $\int \delta(t)\, dt = 1$), the statement $\langle \xi(t)\xi(t') \rangle = 2D\delta(t-t')$ means that ξ has the units $\sqrt{D/t}$. Thus from Eq. (2.3), $x(t)$ must have units of $\sqrt{Dt}$.

The Langevin equation has the great advantage of simplicity. With a bit more work, it is possible to determine higher moments of the position. Furthermore there is a standard prescription to determine the underlying and more fundamental probability distribution of positions. This prescription involves writing a continuum Fokker-Planck equation for the evolution of this probability distribution. The Fokker-Planck equation is in the form of a convection-diffusion equation, namely, the diffusion equation augmented by a term that accounts for a global bias in the stochastic motion. The coefficients in this Fokker-Planck equation are directly related to the parameters in the original Langevin equation. The Fokker-Planck equation can be naturally viewed as the continuum limit of the *master equation*, which represents perhaps the most fundamental way to describe a stochastic process. We will not pursue this conventional approach because we are generally more interested in developing direct approaches to write the master equation.

2.2 Master Equation for the Probability Distribution

Discrete space and time

Consider a random walker on a one-dimensional lattice that hops to the right with probability p or to the left with probability $q = 1 - p$ in a single step. Let $P(x, N)$ be the probability that the particle is at site x at the N^{th} time step. Then evolution of this occupation probability is described by the master equation

$$P(x, N+1) = p\, P(x-1, N) + q\, P(x+1, N). \tag{2.5}$$

Because of translational invariance in both space and time, it is expedient to solve this equation by transform techniques. One strategy is to Fourier transform in space *and* write the generating function (sometimes called the z-transform). Thus multiplying the master equation by $z^{N+1}\, e^{ikx}$ and summing over all N and x gives

$$\sum_{N=0}^{\infty} \sum_{x=-\infty}^{\infty} z^{N+1} e^{ikx} \left[P(x, N+1) = p\, P(x-1, N) + q\, P(x+1, N) \right]. \tag{2.6}$$

We now define the joint transform—the Fourier transform of the generating function

$$P(k, z) = \sum_{N=0}^{\infty} z^N \sum_{x=-\infty}^{\infty} e^{ikx}\, P(x, N).$$

In what follows, either the arguments of a function or the context (when obvious) will be used to distinguish transforms from the function itself. The left-hand side of (2.6) is just the joint transform $P(k, z)$, except that the term $P(x, N = 0)$ is missing. Similarly, on the right-hand side the two factors are just the generating

function at $x-1$ and at $x+1$ times an extra factor of z. The Fourier transform then converts these shifts of ± 1 in the spatial argument to the phase factors $e^{\pm ik}$, respectively. Thus

$$P(k,z) - \sum_{x=-\infty}^{\infty} P(x,N=0)e^{ikx} = zu(k)P(k,z), \tag{2.7}$$

where $u(k) = p\,e^{ik} + q\,e^{-ik}$ is the Fourier transform of the single-step hopping probability. For the initial condition of a particle initially at the origin, $P(x,N=0) = \delta_{x,0}$, the joint transform becomes

$$P(k,z) = \frac{1}{1-zu(k)}. \tag{2.8}$$

We now invert the transform to reconstruct the probability distribution. Expanding $P(k,z)$ in a Taylor series, the Fourier transform of the generating function is simply $P(k,N) = u(k)^N$. Then the inverse Fourier transform is

$$P(x,N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx}\,u(k)^N\,dk, \tag{2.9}$$

To evaluate the integral, we write $u(k)^N = (p\,e^{ik} + q\,e^{-ik})^N$ in a binomial series. This gives

$$P(x,N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx}\sum_{m=0}^{N}\binom{N}{m}p^m\,e^{ikm}\,q^{N-m}\,e^{-ik(N-m)}\,dk. \tag{2.10}$$

The only non-zero term is the one with $m = (N+x)/2$ in which all the phase factors cancel. This leads to the classical binomial probability distribution of a discrete random walk

$$P(x,N) = \frac{N!}{(\frac{N+x}{2})!(\frac{N-x}{2})!}\,p^{\frac{N+x}{2}}\,q^{\frac{N-x}{2}}. \tag{2.11}$$

Finally, using Stirling's approximation, the binomial approaches the Gaussian probability distribution in the long-time limit,

$$P(x,N) \to \frac{1}{\sqrt{2\pi Npq}}e^{-[x-N(p-q)]^2/2Npq}. \tag{2.12}$$

This result is a particular realization of the *central-limit theorem*—namely, that the asymptotic probability distribution of an N-step random walk is *independent* of the form of the single step distribution, as long as the mean displacement $\langle x\rangle$ and the mean-square displacement $\langle x^2\rangle$ in a single step are finite; we will present the central limit theorem in Sec. 2.3.

Continuous time

Alternatively, we can treat the random walk in continuous time by replacing N by continuous time t, the increment $N \to N+1$ with $t \to t+\delta t$, and finally Taylor expanding the master equation (2.5) to first order in δt. These steps give

$$\frac{\partial P(x,t)}{\partial t} = w_+P(x-1,t) + w_-P(x+1,t) - w_0P(x,t) \tag{2.13}$$

where $w_+ = p/\delta t$ and $w_- = q/\delta t$ are the hopping rates to the right and to the left, respectively, and $w_0 = 1/\delta t$ is the total hopping rate from each site. This hopping process satisfies detailed balance, as the total hopping rates *to* a site equal the total hopping rate *from* the same site.

Again, the simple structure of Eq. (2.13) calls out for applying the Fourier transform. After doing so, the master equation becomes

$$\frac{dP(k,t)}{dt} = (w_+e^{ik} + w_-e^{-ik} - w_0)\,P(k,t) \equiv w(k)\,P(k,t). \tag{2.14}$$

For the initial condition $P(x,t=0) = \delta_{x,0}$, the corresponding Fourier transform is $P(k,t=0) = 1$, and the solution to Eq. (2.14) is $P(k,t) = e^{w(k)t}$. To invert this Fourier transform, let's consider the symmetric case

where $w_\pm = 1/2$ and $w_0 = 1$. Then $w(k) = w_0(\cos k - 1)$, and we use the generating function representation for the modified Bessel function of the first kind of order x, $e^{z\cos k} = \sum_{x=-\infty}^{\infty} e^{ikx} I_x(z)$ (10), to give

$$P(k,t) = e^{-t} \sum_{x=-\infty}^{\infty} e^{ikx} I_x(t), \tag{2.15}$$

from which we immediately obtain

$$P(x,t) = e^{-t} I_x(t). \tag{2.16}$$

To determine the probability distribution in the scaling limit where x and t both diverge but x^2/t remains finite, it is more useful to Laplace transform the master equation (2.13) to give

$$sP(x,s) - P(x,t=0) = \frac{1}{2}P(x+1,s) + \frac{1}{2}P(x-1,s) - P(x,s). \tag{2.17}$$

For $x \neq 0$, we solve the resulting difference equation, $P(x,s) = a[P(x+1,s)+P(x-1,s)]$, with $a = 1/2(s+1)$, by assuming the exponential solution $P(x,s) = A\lambda^x$ for $x > 0$; by symmetry $P(x,s) = A\lambda^{-x}$ for $x < 0$. Substituting $P(x,s) = A\lambda^{-x}$ into the recursion for $P(x,s)$ gives a quadratic characteristic equation for λ whose solution is $\lambda_\pm = (1 \pm \sqrt{1-4a^2})/2a$. For all $s > 0$, $\lambda_\pm$ are both real and positive, with $\lambda_+ > 1$ and $\lambda_- < 1$. We reject the solution that grows exponentially with x, thus giving $P_x = A\lambda_-^x$. Finally, we obtain the constant A from the $x = 0$ boundary master equation

$$sP(0,s) - 1 = \frac{1}{2}P(1,s) + \frac{1}{2}P(-1,s) - P(0,s) = P(1,s) - P(0,s). \tag{2.18}$$

The -1 on the left-hand side arises from the initial condition, and the second equality follows by spatial symmetry. Substituting $P(n,s) = A\lambda_-^x$ into Eq. (2.18) gives A, from which we finally obtain

$$P(x,s) = \frac{1}{s+1-\lambda_-}\, \lambda_-^x. \tag{2.19}$$

This Laplace transform diverges at $s = 0$; consequently, we may easily obtain the interesting asymptotic behavior by considering the limiting form of $P(x,s)$ as $s \to 0$. Since $\lambda_- \approx 1 - \sqrt{2s}$ as $s \to 0$, we find

$$P(x,s) \approx \frac{(1-\sqrt{2s})^x}{\sqrt{2s}+s} \sim \frac{e^{-x\sqrt{2s}}}{\sqrt{2s}}. \tag{2.20}$$

We now invert the Laplace transform $P(x,t) = \int_{s_0-i\infty}^{s_0+i\infty} P(x,s)\, e^{st}\, ds$ by using the integration variable $u = \sqrt{s}$. This immediately leads to the Gaussian probability distribution quoted in Eq. (2.26) for the case $\langle x \rangle = 0$ and $\langle x^2 \rangle = 1$.

Continuous space and time

When both space and time are continuous, we expand the master equation (2.5) in a Taylor series to lowest non-vanishing order—second order in space x and first order in time t—we obtain the fundamental *convection-diffusion equation*,

$$\frac{\partial P(x,t)}{\partial t} + v\frac{\partial P(x,t)}{\partial x} = D\,\frac{\partial^2 P(x,t)}{\partial x^2}, \tag{2.21}$$

for the concentration $P(x,t)$. Here $v = (p-q)\delta x/\delta t$ is the bias velocity and $D = \delta x^2/2\delta t$ is the diffusion coefficient. Notice that the factor v/D diverges as $1/\delta x$ in the continuum limit. Therefore the convective term $\frac{\partial P}{\partial x}$ invariably dominates over the diffusion term $\frac{\partial^2 P}{\partial x^2}$. To construct a non-pathological continuum limit, the bias $p-q$ must be proportional to δx as $\delta x \to 0$ so that both the first- and second-order spatial derivative terms are simultaneously finite. For the diffusion equation, we obtain a non-singular continuum limit merely by ensuring that the ratio $\delta x^2/\delta t$ remains finite as both δx and δt approach zero.

To solve the convection-diffusion equation, we introduce the Fourier transform $P(k,t) = \int P(x,t)\, e^{ikx}\, dx$ to simplify the convection-diffusion equation to $\dot{P}(k,t) = (ikv - Dk^2)P(k,t)$, with solution

$$P(k,t) = P(k,0)e^{(ikv-Dk^2)t} = e^{(ikv-Dk^2)t}, \tag{2.22}$$

for the initial condition $P(x, t = 0) = \delta(x)$. We then obtain the probability distribution by inverting the Fourier transform to give, by completing the square in the exponential,

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}}\, e^{-(x-vt)^2/4Dt}. \tag{2.23}$$

Alternatively, we may first Laplace transform in the time domain. For the convection-diffusion equation, this yields the ordinary differential equation

$$sP(x,s) - \delta(x) + vP(x,s) = DP''(x,s), \tag{2.24}$$

where the delta function reflects the initial condition. This equation may be solved separately in the half-spaces $x > 0$ and $x < 0$. In each subdomain Eq. (2.24) reduces to a homogeneous constant-coefficient equation that has exponential solutions. The corresponding solution for the entire line has the form $c_+(x,s) = A_+e^{-\alpha_- x}$ for $x > 0$ and $c_-(x,s) = A_-e^{\alpha_+ x}$ for $x < 0$, where $\alpha_\pm = \left(v \pm \sqrt{v^2 + 4Ds}\right)/2D$ are the roots of the characteristic polynomial. We join these two solutions at the origin by applying the joining conditions of continuity of $P(x,s)$ at $x = 0$, and a discontinuity in $\frac{\partial c}{\partial x}$ at $x = 0$ whose magnitude is determined by integrating Eq. (2.24) over an infinitesimal domain which includes the origin. The continuity condition trivially gives $A_+ = A_- \equiv A$, and the condition for the discontinuity in $P(x,s)$ is $D\left(P'_+|_{x=0} - P'_-|_{x=0}\right) = -1$. This gives $A = 1/\sqrt{v^2 + 4Ds}$. Thus the Laplace transform of the probability distribution is

$$c_\pm(x,s) = \frac{1}{\sqrt{v^2 + 4Ds}}\, e^{-\alpha_\mp |x|}. \tag{2.25}$$

For zero bias, this coincides with Eq. (2.20) and thus recovers the Gaussian probability distribution.

2.3 Central Limit Theorem

The central limit theorem states that the asymptotic $N \to \infty$ probability distribution of an N-step random walk is the universal Gaussian function

$$P(x,N) \to \frac{1}{\sqrt{2\pi N\sigma^2}}\; e^{-(x-\langle x\rangle)^2/2N\sigma^2}, \tag{2.26}$$

where $\langle x\rangle$ and $\langle x^2\rangle$ are respectively the mean and the mean-square displacement for a single step of the walk, and $\sigma^2 = \langle x^2\rangle - \langle x\rangle^2$. A necessary condition for the central limit theorem to hold is that each step of the walk is an independent identically distributed random variable that is drawn from a distribution $p(x)$ such that $\langle x\rangle$ and $\langle x^2\rangle$ are both finite. We now give a simple derivation of this fundamental result. For simplicity we give the derivation for a one-dimensional system, but this derivation can immediately be extended to any dimension.

When the steps of the random walk are independent, the probability distribution after N steps is related to the probability after $N-1$ steps by the recursion (also known as the Chapman-Kolmogorov equation)

$$P_N(x) = \int P_{N-1}(x')p(x' \to x)\, dx'. \tag{2.27}$$

This equation merely states that to reach x in N steps, the walk first reaches an arbitrary point x' in $N-1$ steps and then makes a transition from x' to x with probability $p(x' \to x)$. It is now useful to introduce the Fourier transforms

$$f(k) = \int_{-\infty}^{\infty} f(x)e^{ikx}\, dx \qquad f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(k)e^{-ikx}\, dk$$

to transform Eq. (2.27) to the algebraic equation $P_N(k) = P_{N-1}(k)p(k)$ that we iterate to give $P_N(k) = P_0(k)p(k)^N$. At this stage, there is another mild condition for the central limit theorem to hold—the initial condition cannot be too long range in space. The natural condition is for the random walk to start at the origin, $P_0(x) = \delta x, 0$ for which the Fourier transform of the initial probability distribution is simply $P_0(k) = 1$. Then the Fourier transform of the probability distribution is simply

$$P_N(k) = p(k)^N, \tag{2.28}$$

so that

$$P_N(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} p(k)^N\, e^{-ikx}\, dk. \tag{2.29}$$

To invert the Fourier transform, we now use the fact that the first two moments of $p(x)$ are finite to write the Fourier transform $p(k)$ as

$$\begin{aligned} p(k) &= \int_{-\infty}^{\infty} p(x)\, e^{ikx}\, dx \\ &= \int_{-\infty}^{\infty} p(x)\left[1 + ikx - \frac{1}{2}k^2x^2 + \ldots\right] dx \\ &= 1 + ik\langle x\rangle - \frac{1}{2}k^2\langle x^2\rangle + \ldots \end{aligned}$$

Now the probability distribution is

$$\begin{aligned} P_N(x) &\sim \frac{1}{2\pi}\int_{-\infty}^{\infty} [1 + ik\langle x\rangle - \frac{1}{2}k^2\langle x^2\rangle]^N\, e^{-ikx}\, dx \\ &\sim \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{N\ln[1+ik\langle x\rangle - \frac{1}{2}k^2\langle x^2\rangle]}\, e^{-ikx}\, dx \\ &\sim \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{N[1+ik\langle x\rangle - \frac{k^2}{2}(\langle x^2\rangle - \langle x\rangle^2)]}\, e^{-ikx}\, dx \end{aligned} \tag{2.30}$$

We now complete the square in the exponent and perform the resulting Gaussian integral to arrive at the fundamental result

$$P_N(x) \sim \frac{1}{\sqrt{2\pi N\sigma^2}}\; e^{-(x-N\langle x\rangle)^2/2N\sigma^2}. \tag{2.31}$$

2.4 Connection to First-Passage Properties

An intriguing property of random walks is the transition between *recurrence* and *transience* as a function of the spatial dimension d. Recurrence means that a random walk is certain to return to its starting point; this occurs for $d \leq 2$. Conversely, $d > 2$ the random walk is transient in that there is positive probability for a random walk to never return to its starting point. It is striking that the spatial dimension—and not any other features of a random walk—is the only parameter that determines this transition.

The qualitative explanation for this transition is quite simple. Consider the trajectory of a typical random walk. After a time t a random walk explores a roughly spherical domain of radius $\sqrt{Dt}$ while the total number of sites visited during this walk equals to t. Therefore the density of visited sites within an exploration sphere is $\rho \propto t/t^{d/2} \propto t^{1-d/2}$ in d dimensions. For $d < 2$ this density grows with time; thus a random walk visits each site within the sphere infinitely often and is certain to return to its starting point. On the other hand, for $d > 2$, the density decreases with time and so some points within the exploration sphere never get visited. The case $d = 2$ is more delicate but turns out to be barely recurrent.

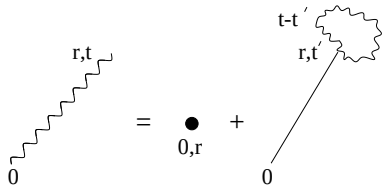


Figure 2.1: Diagrammatic relation between the occupation probability of a random walk (propagation is represented by a wavy line) and the first-passage probability (straight line).

We now present a simple-minded approach to understand this transition between recurrence and transience. Let $P(\mathbf{r},t)$ be probability that a random walk is at $\mathbf{r}$ at time t when it starts at the origin. Similarly, let $F(\mathbf{r},t)$ be the *first-passage probability*, namely, the probability that the random walk visits $\mathbf{r}$ *for the first time* at time t with the same initial condition.

For a random walk to be at $\mathbf{r}$ at time t, the walk must *first* reach $\mathbf{r}$ at some earlier time step t' and then return to $\mathbf{r}$ after $t-t'$ (Fig. 2.1). This connection between $F(\mathbf{r},t)$ and $P(\mathbf{r},t)$ may therefore be expressed as the convolution

$$P(\mathbf{r},t) = \delta_{\mathbf{r},0}\,\delta_{t,0} + \int_0^t F(\mathbf{r},t')\,P(0,t-t')\,dt'. \tag{2.32}$$

The delta function term accounts for the initial condition. The second term accounts for the ways that a walk can be at $\mathbf{r}$ at time t. To reach r at time t, the walk must first reach r at some time $t' \le t$. Once a first passage has occurred, the walk must return to r exactly at time t (and the walk can also return to r at earlier times, so long as the walk is also at r at time t). Because of the possibility of multiple visits to $\mathbf{r}$ between time t' and t, the return factor involves P rather than F. This convolution equation is most conveniently solved in terms of the Laplace transform to give $P(\mathbf{r},s) = \delta_{\mathbf{r},0} + F(\mathbf{r},s)P(0,s)$. Thus we obtain the fundamental connection

$$F(\mathbf{r},s) = \begin{cases} \dfrac{P(\mathbf{r},s)}{P(0,s)}, & \mathbf{r} \neq 0 \\ 1 - \dfrac{1}{P(0,s)}, & \mathbf{r} = 0, \end{cases} \tag{2.33}$$

in which the Laplace transform of the first-passage probability is determined by the corresponding transform of the probability distribution of diffusion $P(\mathbf{r},t)$.

We now use the techniques of Section A.2 to determine the time dependence of the first-passage probability in terms of the Laplace transform for the occupation probability. For isotropic diffusion, $P(\mathbf{r}=0,t) = (4\pi Dt)^{-d/2}$ in d dimensions and the Laplace transform is $P(0,s) = \int_0^\infty P(0,t)\,e^{-st}\,dt$. As discussed in Section A.2, this integral has two fundamentally different behaviors, depending on whether $\int^\infty P(0,t)\,dt$ diverges or converges. In the former case, we apply the last step in Eq. (A.6) to obtain

$$P(0,s) \propto \int^{t^*=1/s} (4\pi Dt)^{-d/2}\,dt \sim \begin{cases} \mathcal{A}_d (t^*)^{1-d/2} = \mathcal{A}_d s^{d/2-1}, & d<2 \\ \mathcal{A}_2 \ln t^* = -\mathcal{A}_2 \ln s, & d=2, \end{cases} \tag{2.34}$$

where the dimension-dependent prefactor $\mathcal{A}_d$ is of the order of 1 and does not play any role in the asymptotic behavior.

For $d>2$, the integral $\int^\infty P(0,t)\,dt$ converges and one has to be more careful to extract the asymptotic behavior by studying $P(0,1) - P(0,s)$. By such an approach, it is possible to show that $P(0,s)$ has the asymptotic behavior

$$P(0,s) \sim (1-\mathcal{R})^{-1} + B_d s^{d/2-1} + \ldots, \qquad d>2, \tag{2.35}$$

where $\mathcal{R}$ is the *eventual return* probability, namely, the probability that a diffusing particle random walk ultimately reaches the origin, and B_d is another dimension-dependent constant of the order of 1. Using these results in Eq. (2.33), we infer that the Laplace transform for the first-passage probability has the asymptotic behaviors

$$F(0,s) \sim \begin{cases} 1 - A_d s^{1-d/2}, & d<2 \\ 1 + A_2 (\ln s)^{-1}, & d=2 \\ \mathcal{R} + B_d (1-\mathcal{R})^2 s^{d/2-1}, & d>2, \end{cases} \tag{2.36}$$

From this Laplace transform, we determine the time dependence of the survival probability by approximation (A.9); that is,

$$F(0, s = 1 - 1/t^*) \sim \int_0^{t^*} F(0,t)\,dt \equiv T(t^*), \tag{2.37}$$

where $T(t)$ is the probability that the particle gets trapped (reaches the origin) by time t. For what follows, we also define the survival probability $S(t) = 1 - T(t)$, which is simply the probability that the particle has

not reached the origin by time t. Here the trick of replacing an exponential cutoff by a sharp cutoff provides an extremely easy way to invert the Laplace transform. From Eqs. (2.36) and (2.37) we thus find

$$S(t) \sim \begin{cases} A_d t^{d/2-1}, & d<2 \\ A_2 (\ln t)^{-1}, & d=2 \\ (1-\mathcal{R}) + C_d\,(1-\mathcal{R})^2\, t^{1-d/2}, & d>2. \end{cases} \tag{2.38}$$

where C_d is another d-dependent constant of the order of 1. Finally, the time dependence of the first-passage probability may be obtained from the basic relation $1 - S(t) \sim \int^t F(0,t)\,dt$ to give

$$F(0,t) = -\frac{\partial S(t)}{\partial t} \propto \begin{cases} t^{d/2-2,} & d<2 \\ t^{-1}(\ln t)^{-2}, & d=2 \\ t^{-d/2}, & d>2. \end{cases} \tag{2.39}$$

It is worth emphasizing several important physical ramifications of the above first-passage properties. First, the asymptotic behavior is determined by the spatial dimension only and that there is a dramatic change in behavior when $d = 2$. For $d \leq 2$, the survival probability $S(t)$ ultimately decays to zero. This means that a random walk is *recurrent* and is certain to eventually return to its starting point, and indeed visit *any* site of an infinite lattice. Finally, because a random walk has no memory, it is "renewed" every time a specific lattice site is reached. Thus recurrence also implies that every lattice site is visited infinitely often.

We can give is a simple physical explanation for this efficient visitation of sites. After a time t a random walk explores a roughly spherical domain of radius $\sqrt{Dt}$. The total number of sites visited during this exploration is also proportional to t. Consequently in d dimensions, the density of visited sites within this exploration sphere is $\rho \propto t/t^{d/2} \propto t^{1-d/2}$. For $d < 2$, ρ diverges as $t \to \infty$ and a random walk visits each site within the sphere infinitely often. This feature is termed *compact exploration*. Paradoxically, although every site is visited with certainty, these visitations take forever because the mean time to return to the origin, $\langle t \rangle = \int t\,F(0,t)\,dt$, diverges for all $d \leq 2$.

Finally, we outline a useful technique to compute *where* on a boundary is a diffusing particle absorbed and *when* does this absorption occur. This method will provide helpful in understanding finite-size effect in reaction kinetics. For simplicity, consider a symmetric nearest-neighbor random walk in the finite interval $[0, 1]$. Let $\mathcal{E}_+(x)$ be the probability that a particle, which starts at x, eventually hits $x = 1$ *without* hitting $x = 0$. This eventual hitting probability $\mathcal{E}_+(x)$ is obtained by summing the probabilities for all paths that start at x and reach 1 without touching 0. Thus

$$\mathcal{E}_+(x) = \sum_p \mathcal{P}_p(x), \tag{2.40}$$

where $\mathcal{P}_p(x)$ denotes the probability of a path from x to 1 that does not touch 0. The sum over all such paths can be decomposed into the outcome after one step (the factors of 1/2 below) and the sum over all path remainders from the location after one step to 1. This gives

$$\mathcal{E}_+(x) = \sum_p \left[\frac{1}{2}\,\mathcal{P}_p(x+\delta x) + \frac{1}{2}\,\mathcal{P}_p(x-\delta x)\right] = \frac{1}{2}[\mathcal{E}_+(x+\delta x) + \mathcal{E}_+(x-\delta x)]. \tag{2.41}$$

By a simple rearrangement, this equation is equivalent to

$$\Delta^{(2)}\mathcal{E}_+(x) = 0, \tag{2.42}$$

where $\Delta^{(2)}$ is the second-difference operator. Notice the opposite sense of this recursion formula compared to the master equation Eq. (2.5) for the probability distribution. Here $\mathcal{E}_+(x)$ is expressed in terms of *output from* x, while in the master equation, the occupation probability at x is expressed in terms of *input to* x. For this reason, Eq. (2.41) is sometimes referred to as a backward master equation. This backward equation is just the Laplace equation and gives a hint of the deep relation between first-passage properties, such as the exit probability, and electrostatics. Equation (2.42) is subject to the boundary conditions $\mathcal{E}_+(0) = 0$

and $\mathcal{E}_+(1) = 1$; namely if the walk starts at 1 it surely exits at 1 and if the walk starts at 0 it has no chance to exit at 1. In the continuum limit, Eq. (2.42) becomes the Laplace equation $\mathcal{E}'' = 0$, subject to appropriate boundary conditions. We can now transcribe well-known results from electrostatics to solve the exit probability. For the one dimensional interval, the result is remarkably simple: $\mathcal{E}_+(x) = x$!

This exit probability also represents the solution to the classic "gambler's ruin" problem: let x represent your wealth that changes by a small amount dx with equal probability in a single bet with a Casino. You continue to bet as long as you have money. You lose if your wealth hits zero, while you break the Casino if your wealth reaches 1. The exit probability to $x = 1$ is the same as the probability that you break the Casino.

Let's now determine the mean time for a random walk to exit a domain. We focus on the *unconditional* exit time, namely, the time for a particle to reach *any* point on the absorbing boundary of this domain. For the symmetric random walk, let the time increment between successive steps be δt, and let $t(x)$ denote the average exit time from the interval $[0, 1]$ when a particle starts at x. The exit time is simply the time for each exit path times the probability of the path, averaged over all trajectories, and leads to the analog of Eq. (2.40)

$$t(x) = \sum_p \mathcal{P}_p(x)\, t_p(x), \tag{2.43}$$

where $t_p(x)$ is the exit time of a specific path to the boundary that starts at x.

In analogy with Eq. (2.41), this mean exit time obeys the recursion

$$t(x) = \frac{1}{2}\left[(t(x+\delta x) + \delta t) + (t(x-\delta x) + \delta t)\right], \tag{2.44}$$

This recursion expresses the mean exit time starting at x in terms of the outcome one step in the future, for which the initial walk can be viewed as restarting at either $x+\delta x$ or $x-\delta x$, each with probability 1/2, but also with the time incremented by δt. This equation is subject to the boundary conditions $t(0) = t(1) = 0$; the exit time equals zero if the particle starts at the boundary. In the continuum limit, this recursion formula reduces to the Poisson equation $Dt''(x) = -1$. For diffusion in a d-dimensional domain with absorption on a boundary B, the corresponding Poisson equation for the exit time is $D\nabla^2 t(\mathbf{r}) = -1$, subject to the boundary condition $t(\mathbf{r}) = 0$ for $\mathbf{r} \in B$. Thus the determination of the mean exit time has been recast as a time-independent electrostatic problem! For the example of the unit interval, the solution to the Laplace equation is just a second-order polynomial in x. Imposing the boundary conditions immediately leads to the classic result

$$t(x) = \frac{1}{2D}\; x(1-x). \tag{2.45}$$

First passage probability and the gambler's ruin problem

Consider a random walk in a finite interval of length N. The two boundary sites are absorbing, *i.e.*, the random walker immediately disappears upon reaching these sites. Suppose that the starting position of the random walk is n, with $0 \le n \le N$. What is F_n, the probability that the walker first reaches the boundary at site N? We can write a simple recursion formula for the first-passage probability. With probability 1/2, the walk steps to site $n-1$, at which point the exit probability to site N is F_{n-1}. Similarly, the walk steps to site $n+1$ with probability 1/2, where the exit probability is F_{n+1}. Thus the first passage probability satisfies the discrete Poisson equation

$$F_n = \frac{1}{2}(F_{n-1} + F_{n+1}), \tag{2.46}$$

with the boundary conditions $F_0 = 0$ and $F_N = 1$. The solution is simple:

$$F_n = \frac{n}{N}. \tag{2.47}$$

This first passage probability also solves a neat probability theory problem. In a fair coin-toss game, the probability that a gambler ruins a Casino equals the wealth of the gambler divided by the combined wealth of the gambler and casino. Gambling is most definitely a bad idea...

2.5 The Reaction Rate

Suppose that you wanted to hit the side of a barn using an ensemble of blind riflemen that fire bullets in random directions as your incident beam. What is the rate at which the barn is hit? Theorists that we are, let's model the barn as a sphere of radius R. A patently obvious fact is that if the radius of the barn is increased, the number of bullets that hit our theoretical barn increases as its cross-sectional area. In d spatial dimensions, the cross section therefore scales as R^{d-1}. Now suppose that we take away the rifles from our blind marksmen and give them the task of hitting the barn simply by wandering around. Surprisingly, the rate at which the blind riflemen diffuse to the barn is proportional to R^{d-2} for $d > 2$. Thus in the physical case of 3 dimensions, the absorption rate is proportional to the sphere radius rather than to its cross section! Even more striking— for $d \leq 2$ the absorption rate no longer depends on the radius of the absorbing sphere. The rate at which diffusing particles hit an absorbing sphere is the underlying mechanism of diffusion-controlled reactions. Because of the centrality of this topic to reaction kinetics and because it represents a nice application of first-passage ideas, we now determine this reaction rate.

As in the original Smoluchowski theory for the reaction rate, we fix a spherical absorbing particle of mass m_i radius R_i at the origin, while a gas of non-interacting particles each of mass m_j and radii R_j freely diffuses outside the sphere. The separation between the absorbing sphere and a background particle diffuses with diffusion coefficient $D_i + D_j$, where D_i is the diffusion coefficient of a droplet of radius R_i. When the separation first reaches $a = R_i + R_j$, reaction occurs. The reaction rate is then identified as the flux to an absorbing sphere of radius a by an effective particle with diffusivity $D = D_i + D_j$.

The concentration of background particles around the absorbing sphere thus obeys the diffusion equation

$$\frac{\partial c(\vec{r}, t)}{\partial t} = D\nabla^2 c(\vec{r}, t), \tag{2.48}$$

subject to the initial condition $c(\vec{r}, t = 0) = 1$ for $r > a$ and the boundary conditions $c(r = a, t) = 0$ and $c(r \to \infty, t) = 1$. The reaction rate is then identified with the integral of the flux over the sphere surface

$$K(t) = -D \int_S \Big| \frac{\partial c(\vec{r}, t)}{\partial r} \Big|_{r=a} d\Omega. \tag{2.49}$$

There are two regimes of behavior as a function of the spatial dimension. For $d > 2$, the loss of reactants at the absorbing sphere is sufficiently slow that it is replenished by the re-supply from larger distances. A steady state is thus reached and the reaction rate K is finite. In this case, the reaction rate can be determined more simply by solving the time-independent Laplace equation, rather than the diffusion equation (2.48).

The solution to the Laplace equation with the above initial and boundary conditions is

$$c(r) = 1 - \left(\frac{a}{r}\right)^{d-2}.$$

The flux is then $-D\frac{\partial c}{\partial r}|_{r=a} = D(d-2)/a$ and the total current is the integral of this flux over the surface of the sphere $K = (d-2)\Omega_d\, Da^{d-2}$, where $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the area of a unit sphere in d dimensions. We translate this flux into the reaction kernel for aggregation by expressing a and D in terms of the parameters of the constituent reactants to give

$$K_{ij} = (d-2)\Omega_d\, (D_i + D_j)(R_i + R_j)^{d-2}. \tag{2.50}$$

We can express this result as a function of reactant masses only for the physical case of three dimension by using $R_i \propto i^{1/3}$, while for the diffusion coefficient, we use the Einstein-Stokes relation $D_i = kT/(6\pi\eta R_i) \propto i^{-1/3}$, where kT is the thermal energy and η is the viscosity coefficient to obtain

$$K_{ij} \propto \frac{2kT}{3\eta}\, (R_i^{-1} + R_j^{-1})(R_i + R_j). \tag{2.51}$$

What happens for $d < 2$? We could solve the diffusion equation with the absorbing boundary condition and the unit initial condition, from which the time-dependent flux and thereby a time-dependent reaction rate

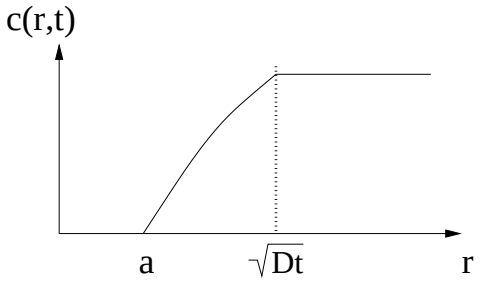


Figure 2.2: Sketch of the concentration about an absorbing sphere according to the quasi-static approximation. The near- and far-zone concentrations match at $r = \sqrt{Dt}$.

can be deduced. However, it is simpler and more revealing to apply the general *quasi-static* approximation. Because of its simplicity and general utility, we now present the quasi-static calculation of the reaction rate. The basis of the quasi-static approximation is that the region exterior to the absorbing sphere naturally divides into "near" and "far" zones. In the near zone, which extends to a distance $\sqrt{Dt}$ from the sphere, diffusing particles have ample time to explore this near zone thoroughly and the concentration is nearly time independent. In the complementary far zone there is negligible depletion because diffusing particles that are more distant than $\sqrt{Dt}$ typically will not hit the sphere in a time t. Thus in the far zone the concentration $c(r) = 1$ for $r > \sqrt{Dt}$.

Based on this picture, we merely solve the Laplace equation in the near zone $a < r < \sqrt{Dt}$ with the *time-dependent* boundary conditions $c(r = \sqrt{Dt}) = 1$, to match to the static far-zone solution, and $c(a) = 0$. The general solution is $c(r) = A + Br^{2-d}$, and matching to the boundary conditions gives

$$c(r,t) = \frac{1-(a/r)^{d-2}}{1-(a/\sqrt{Dt})^{d-2}} \to \left(\frac{\sqrt{Dt}}{r}\right)^{d-2} \qquad t\to\infty \qquad \text{for } d=1. \tag{2.52a}$$

For $d = 2$, we can still apply the same quasi-static approach because diffusion is still recurrent, so that a qualitatively similar depletion layer builds up around the absorbing sphere. Now, however, the general solution to the Laplace equation is $c(r) = A + B\ln r$. Apply the boundary conditions at $r = a$ and $r = \sqrt{Dt}$ leads to

$$c(r,t) = \frac{\ln(r/a)}{\ln(\sqrt{Dt}/a)} \to \frac{\ln r}{\ln t} \qquad t\to\infty \qquad \text{for } d=2. \tag{2.52b}$$

Finally, we substitute the above expressions for the concentration into the definition of the time-dependent reaction rate from Eq. (2.49) to obtain the reaction rate.

$$K(t) \propto \begin{cases} D\times(Dt)^{(d-2)/2} & d<2; \\ \dfrac{4\pi D}{\ln\left(Dt/a^2\right)} & d=2; \\ Da^{d-2} & d>2. \end{cases} \tag{2.53}$$

Notice that the rate does not depend on the cluster radius for $d \le 2$. This surprising fact arises because of the recurrence of diffusion in $d \le 2$ so that two diffusing particles are guaranteed to eventually meet independent of their radii.

Problems

Section 2.2

1. Find the generating function for the Fibonacci sequence, $F_n = F_{n-1} + F_{n-2}$, with the initial condition $F_0 = F_1 = 1$; that is, determine $F(z) = \sum_0^\infty F_n z^n$. Invert the generating function to find a closed form expression for F_n.

2. Consider a random walk in one dimension in which a step to the right of length 2 occurs with probability 1/3 and a step to the left of length 1 occurs with probability 2/3. Investigate the corrections to the isotropic Gaussian that characterizes the probability distribution in the long-time limit. *Hint:* Consider the behavior of moments beyond second order, $\langle x^k \rangle$ with $k > 2$.

3. Solve the gambler's ruin problem when the probability of winning in a single bet is p. The betting game is repeated until either you are broke or the casino is broken. Take the total amount of capital to be $\$N$ and you start with $\$n$. What is the probability that you will break the casino? Also determine the mean time until the betting is over (either you are broke or the Casino is broken). *More advanced:* Determine the mean time until betting is over with the condition that: (i) you are broke, and (ii) you break the Casino. Solve this problem both for fair betting and biased betting.

4. Consider the gambler's ruin problem under the assumptions that you win each bet with probability $p \neq 1/2$, but that the casino has an infinite reserve of money. What is the probability that you break the casino as a function of p? For those values of p where you break the casino, what is the average time for this event to occur?

Section 2.4

5. For $\mathbf{r} \neq 0$ and $t > 0$, explicitly verify Eq. (2.32) in one dimension.

   Solution.

Notes

The field of random walks, diffusion, and first-passage processes are classic areas of applied probability theory and there is a corresponding large literature. For the more probabilistic aspects of random walks and probability theory in general, we recommend 15; 3; 11; 9. For the theory of random walks and diffusion from a physicist's perspective, we recommend 12; 13; 14. For first-passage properties, please consult 8; 9.

Chapter 3

COLLISIONS

3.1 Background

The foundational Boltzmann transport equation (BTE), which describes how a gas evolves by molecular collisions, normally appears early in a non-equilibrium statistical physics course. Some basic consequences of the BTE include the Maxwell-Boltzmann velocity distribution, the determination of transport coefficients, and the derivation of the Navier-Stokes equations of hydrodynamics. The latter two items involve the formidable and subtle Chapman-Enskog expansion, and the complexity of this approach makes kinetic theory an intimidating subject with which to begin studying non-equilibrium statistical physics. We think it is more useful pedagogically to focus on simpler and explicitly solvable models, such as the Lorentz gas, where a test particle interacts with a fixed scattering background, and the Maxwell model, where the collision rate is independent of the velocities of collision partners. The solutions of these models help illustrate the master equation methodology that we use throughout this book, as well as provide intuition about how to deal with more realistic collisional dynamics.

The Maxwell-Boltzmann Distribution

As a preliminary, let's derive the Maxwell-Boltzmann (MB) velocity distribution for a classical gas of identical molecules. The gas is in equilibrium at temperature T and two molecules scatter elastically when they are sufficiently close due to a short-range repulsive intermolecular potential. Let $P(\mathbf{v})\, d\mathbf{v}$ be the probability to find a molecule within a range $d\mathbf{v}$ about $\mathbf{v}$ when the temperature is T. The MB distribution is based on two fundamental assumptions:

1. Spatial symmetry, which implies that the MB distribution is isotropic:

$$P(\mathbf{v}) = P(v^2); \tag{3.1}$$

   i.e., the distribution depends only on the magnitude v and not on the direction of $\mathbf{v}$.

2. Molecular chaos, which implies that different velocity components are uncorrelated. Together with symmetry, this assumption allows the velocity distribution to be factorized as

$$P(v^2) = p(v_x^2)\, p(v_y^2)\, p(v_z^2), \tag{3.2}$$

   where $p(v_i^2, T)\, dv_i$ is the probability that the i^{th} velocity component is in a range dv_i about v_i.

Let us now derive the MB distribution using Maxwell's original approach. First, take the logarithm of (3.2) and then differentiate with respect to one velocity component to give

$$\frac{\partial \ln P(v^2)}{\partial v_i^2} = \frac{\partial \ln P}{\partial v^2}\frac{\partial v^2}{\partial v_i^2} = \frac{\partial \ln P}{\partial v^2} = \frac{\partial \ln p(v_i^2)}{\partial v_i^2}\ .$$

Since this equation holds for any component, the right-hand side must be constant. Thus we conclude that $\ln P = a + bv^2$. The unknown constants in $P = A\, e^{bv^2}$ can then be found by normalization and equipartition (each degree of freedom contributes an average energy per particle[1] of $\frac{1}{2}m\langle v_x^2\rangle = \frac{1}{2}T$) to give

$$P(\mathbf{v}) = \left(\frac{m}{2\pi T}\right)^{3/2} e^{-m\mathbf{v}^2/2T} \ . \tag{3.3}$$

Notice that almost nothing about the interparticle potential enters into this argument. We only require that there exists some scattering mechanism that conserves mass, momentum, and energy in each collision, so that a steady state actually exists. With these modest requirements, the steady-state velocity distribution is the Maxwell-Boltzmann form, independent of the intermolecular interaction. After this derivation, a cynical physics student might ask: isn't it simpler still to "derive" by MB distribution from the Boltzmann factor, $e^{-E/T}$ of equilibrium statistical mechanics? While it is true that there is no work involved in writing $P \propto e^{-E/T}$, considerable effort is involved in developing the entire apparatus of equilibrium statistical mechanics, upon which the Boltzmann factor is based.

The Boltzmann Transport Equation (BTE)

It is instructive present some of the basic features of the BTE as a prelude for discussing the collisional processes of this chapter. In classical kinetic theory, the fundamental quantity is the space- and velocity-dependent distribution, $P(\mathbf{r},\mathbf{v};t)\, d\mathbf{r}\, d\mathbf{v}$, defined as the probability that a particle is within a range $d\mathbf{r}\, d\mathbf{v}$ about $\mathbf{r},\mathbf{v}$ at temperature T. This distribution evolves according to the BTE, which we write in the slightly symbolic form:

$$\left(\frac{\partial}{\partial t} + v_i\frac{\partial}{\partial r_i} + F_i\frac{\partial}{\partial v_i}\right) P(\mathbf{r},\mathbf{v};t) = \\ = \int |\mathbf{v}-\mathbf{v}'|\, \frac{d\sigma}{d\Omega}\, \left[P_2(\mathbf{r},\mathbf{v};\mathbf{r},\mathbf{v}') - P_2(\mathbf{r},\mathbf{v}'';\mathbf{r},\mathbf{v}''')\right] \delta(\mathbf{p}^{(i)} - \mathbf{p}^{(f)})\delta(E^{(i)} - E^{(f)}) d\mathbf{v}'\, d\mathbf{v}'\, d\mathbf{v}''' \tag{3.4}$$

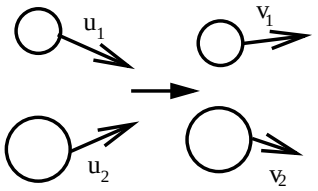


Figure 3.1: Binary collision with pre-collision velocities u_1 and u_2 and post-collision velocities v_1 and v_2.

The terms on the left-hand side represent the change in $P(\mathbf{r},\mathbf{v};t)$ due to particle motion in the absence of collisions; these terms simply comprise the total time derivative. We use the Einstein summation convention that repeated Cartesian indices are summed. The right-hand side represents the change in $P(\mathbf{r},\mathbf{v};t)$ due to collisions. The second term represents the loss of particles with phase space coordinates $\mathbf{r},\mathbf{v}$ due to collisions with particles whose coordinates are $\mathbf{r},\mathbf{v}'$. The total collision rate involves the product of the relative speeds of the two particles, the cross section for their scattering into any outgoing state, and the probability that the collision partners are at the same spatial point. This latter probability is expressed by the *two-body* correlation function P_2. The first term represents a gain in which particles with coordinates $\mathbf{r},\mathbf{v}''$ and $\mathbf{r},\mathbf{v}'''$ collide, with one of the outgoing particles having coordinates $\mathbf{r},\mathbf{v}$ and the other $\mathbf{r},\mathbf{v}'$ By time reversal invariance, both the collision cross-section and the relative velocity of the gain term are the same as that of the loss term. Finally, the delta functions impose overall momentum and energy conservation; here the superscripts denote the total momentum and energy in the initial and final states.

As written, the BTE is not closed: the equation of motion of the first single-particle distribution involves a two-body distribution. However, the equation for the two-body distribution involves the three-body

[1] In this chapter, we set Boltzmann's constant $k_B = 1$

distribution, etc. To make this infinite equation hierarchy closed, it is necessary to break the hierarchy at some level by replacing n-body distributions as products of lower-order distributions. For the BTE, the conventional approach is to replace two-body distributions by the product of single-body distributions; that is,

$$P_2(\mathbf{r},\mathbf{v};\mathbf{r},\mathbf{v}') \to P(\mathbf{r},\mathbf{v})P(\mathbf{r},\mathbf{v}').$$

This is the *molecular chaos* assumption, which is crucial for making further progress in solving the BTE.

At this point, it is worthwhile to highlight the assumptions underlying the BTE and some of their implications:

1. The gas is sufficiently dilute that the one-body distribution $P(\mathbf{r},\mathbf{v};t)$ describes the state of the gas accurately.

2. The gas is sufficiently dilute that only binary collisions are important.

3. Particles move with constant velocity between collisions and undergo scattering events according to classical mechanics when two molecules are in physical contact. We can therefore view molecules as *hard spheres.*

4. We decompose a two-body distribution as a product of one-body distributions; that is, the states of two particles are uncorrelated. This is the *mean-field*, or the *molecular chaos* assumption. This decomposition may be regarded as a truncation at second order of an infinite hierarchy of equations that describe the evolution of n-body correlation functions in terms of $(n+1)$-body correlations. For the BTE, the decomposition of the two-particle correlation function also leads to the crucial feature of the breaking of time-reversal symmetry.

The solution of the equation hierarchy for multi-particle correlation functions has not been obtained for any finite level of truncation, even for the simplest second-order truncation of the BTE. To make progress for this second-order truncation, one normally resorts to the perturbative Chapman-Enskog expansion in which the small parameter is the ratio of the mean-free path to a characteristic system size and the distribution is expressed as a local Maxwell-Boltzmann distribution plus correction terms.

It is at this stage that we terminate our discussion of the BTE because of its technical complications and instead turn to idealized and simpler descriptions of kinetic theory to avoid the daunting calculations needed to deal the BTE. While these idealizations contain some element of fantasy, they are still sufficiently grounded in the underlying kinetics that they provide useful insights about physical reality. An especially useful simplification in this spirit is the *Maxwell model*. The basic feature of this model is the replace the relative velocity inside the collision integral by a constant, a device that greatly facilitates further analysis.

This seemingly drastic approximation can also be justified physically, as a collision rate that is independent of the relative velocity arises for a specific form of the interaction potential between molecules. Suppose that this potential has the form $U(r) = \frac{A}{r^n}$. We may then estimate the closest approach of two molecules in equilibrium at temperature T by equating their interaction energy with the kinetic energy, $\frac{A}{r^n} = \frac{1}{2}mv^2 = T$. Thus $r \propto v^{-2/n}$; we may view this distance as the scattering radius of each molecule. Then the overall collision rate has the following dependence on the relative velocity:

$$u\,\frac{d\sigma}{d\Omega} \propto ur^{d-1} \sim u^{1-2(d-1)/n}\ .$$

Thus the collision rate is *independent* of the relative velocity when $n = 1/[2(d-1)]$. In the physical case of $d = 3$, the Maxwell model corresponds to an r^{-4} interaction potential between molecules.

Another description is that of *very hard particles* in which the overall collision rate is written as the *square* of the relative velocity between two molecules. Because of this quadratic interaction, the interaction is "harder" than the hard-sphere interaction of classical kinetic theory. On the other hand, the Maxwell model is a "softer" interaction because of the velocity-independent interaction. In fact, the Maxwell model and very hard particles can be viewed bounds on classical kinetic theory.

3.2 The Lorentz Gas

3.3 Inelastic Gases

While kinetic theory traditionally deals with gases whose molecules interact via elastic collisions, many new and unexpected phenomena arise when collisions are *inelastic*. Perhaps the most spectacular is the large-scale clustering and the *inelastic collapse* of a freely-evolving inelastic gas. At a more microscopic but still fundamental level, the velocity distribution of inelastic gases is generally not Gaussian. We now discuss these and related basic kinetic features of inelastic gases.

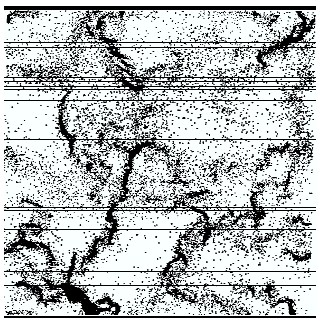


Figure 3.2: Example of inelstic clustering. The resitution coefficient is 0.6 and there are on average 500 collisions per particle. The number of particles is 40000 and the area fraction is 0.05.

where to put the statement of the origin of non-Gaussianity

Haff's Law

Because of inelastic collisions, the mean kinetic energy, and correspondingly, the temperature decreases with time. What is the time dependence of this cooling? Using simple ideas from kinetic theory, we can determine this time dependence under the assumption that the gas remains spatially homogeneous. As we shall see, homogeneity is a reasonable approximation at early times, but this naive assumption is quite wrong in the long-time limit because large-scale density heterogeneities arise.

We estimate the cooling rate of an inelastic gas from the outcome of a typical collision. The kinetic energy lost in an inelastic collision is $\Delta T = -\epsilon(1-\epsilon)(\Delta v)^2 \approx -\epsilon(\Delta v)^2$, with Δv the relative velocity between colliding particles. We consider the quasi-elastic limit $\epsilon \to 0$, where the spatial homogeneity assumption is a reasonable approximation over a non-negligible time range. The typical time Δt between collisions is roughly $\ell/\Delta v$, with ℓ the mean-free path. We assume that a single scale characterizes all velocities so that $\Delta v \sim v \sim \sqrt{T}$. Putting these elements together, the temperature cooling rate therefore is $\frac{dT}{dt} \approx -\frac{\Delta T}{\Delta t} \propto -\epsilon T^{3/2}$. From this rate equation, the temperature decays as

$$T(t) = T_0(1 + c\epsilon t)^{-2}, \tag{3.5}$$

with the constant c of the order of one. Notice that the gas remains effectively elastic $T(t) \approx T(0)$ for $t \ll \epsilon^{-1}$. This time range can be substantial if the dissipation by collisions is sufficiently small. Beyond this time range, the temperature decays algebraically in time, $T(t) \sim (\epsilon t)^{-2}$.

Inelastic Collapse in One Dimension

In one dimension, many aspects of inelastic collapse can be quite simply understood and it is therefore instructive to first focus on this case. In one dimension, the post-collision velocities (v_1, v_2) are related to

the pre-collision velocities (u_1, u_2) of two approaching equal-mass particles by momentum conservation:

$$v_1 = \frac{1}{2}(1-r)u_1 + \frac{1}{2}(1+r)u_2 \equiv \epsilon u_1 + (1-\epsilon)u_2 \qquad (3.6)$$
$$v_2 = \frac{1}{2}(1+r)u_1 + \frac{1}{2}(1+r)u_2 \equiv (1-\epsilon)u_1 + \epsilon u_2,$$

where $0 \le r \le 1$ is the restitution coefficient and, for later convenience, we define the collision parameter $\epsilon \equiv \frac{1}{2}(1-r)$. The restitution coefficient is defined by the incoming momenta $\pm p$ in the center-of-mass reference frame becoming $\mp rp$ after the collision. Setting the particle mass to one, the energy loss in a collision is

$$\Delta E = (u_1^2 + u_2^2)/2 - (v_1^2 + v_2^2)/2 = -\epsilon(1-\epsilon)(u_1 - u_2)^2.$$

The energy loss is maximal for a completely inelastic collision ($r = 0$, $\epsilon = 1/2$) and vanishes for elastic collisions ($r = 1$, $\epsilon = 0$).

Consider now N particles that are initially at $x_i = i$, $i = 2, 3, \ldots, N$. Particle 1 is incident on particle 2 with velocity $+1$, leading to a subsequent series of inelastic collisions according to (3.6). How does this deterministic system evolve with time?

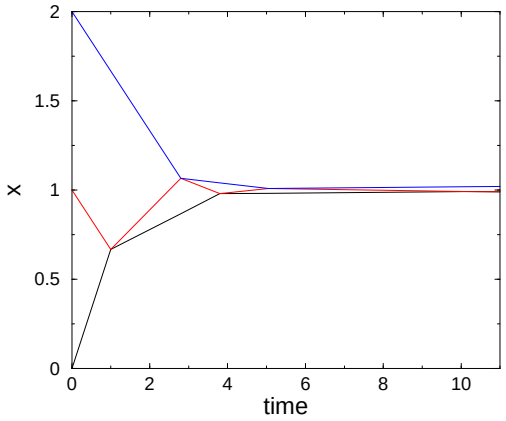


Figure 3.3: Illustration of the collision sequence for 3 unit-mass particles in one dimension. Particle 1 initially moves at velocity $+1$, while particles 2 & 3 are at rest. Shown are the particle worldlines in the center-of-mass reference frame that moves at velocity $+1/3$ for restitution coefficient $r = 0.11 > r_c(3)$. The trajectories initially converge, but after the last collision between particles 1 & 2 at $t \approx 10.164$, the trajectories diverge.

When there are two particles, they collide at most once. The first non-trivial case is $N = 3$, where a sequence of alternating 12 and 23 collisions occurs (Fig. 3.3). Without loss of generality, let us assume that the first collision is 12. The total number of collisions and the ultimate fate of the particles — either collapse or diverging — depends on the restitution coefficient. It is convenient to represent the collision outcome in matrix notation. Then after the pair of collisions 12 and 23, the post-collision velocities v_i are related to the pre-collision velocities u_i by Eq. (3.6),

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = M_{23}M_{12}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \qquad \text{with} \qquad M_{12} = \begin{pmatrix} \epsilon & 1-\epsilon & 0 \\ 1-\epsilon & \epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \epsilon & 1-\epsilon \\ 0 & 1-\epsilon & \epsilon \end{pmatrix}. \qquad (3.7)$$

After $2n$ collisions, the particle velocities are given by $\mathbf{v} = M^n\mathbf{u}_0$, where $M = M_{23}M_{12}$ and $\mathbf{u}_0$ is the initial velocity vector, while after $2n+1$ collisions, the velocities are given by $\mathbf{v} = M_{12}M^n\mathbf{u}_0$. The nature of the collisions follows from the eigenvalues of the matrix M whose values are

$$\lambda_{1,2} = \frac{1}{2}\left[(\epsilon^2 + 2\epsilon - 1) \pm \sqrt{(1-2\epsilon-\epsilon^2)^2 - 4(2\epsilon-1)^2}\right], \quad \lambda_3 = 1.$$

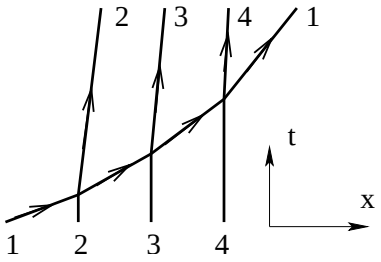


Figure 3.4: "Bending" of a test particle as it collides inelastically to penetrate a static particle array. Particle labels are exchanged in each collision so that particle 1 is decelerated by collisions.

When $\epsilon > \epsilon_c = 2\sqrt{3}-3$, corresponding to restitution coefficient $r < r_c = 7-4\sqrt{3}$, then $\lambda_{1,2}$ are both real with absolute values less than 1. Consequently the particle velocities asymptotically decay as $[\max(|\lambda_1|, |\lambda_2|)]^n$. Conversely for $r > r_c$, the collision sequence terminates after a finite number as the particles ultimate diverge.

When $N > 3$, the collision sequence is not necessarily periodic and the matrix formulation no longer provides the exact value of the critical restitution coefficient for collapse. However, we can give a simple heuristic argument for the dependence of the critical restitution coefficient r_c on N for large N. Consider a test particle with speed 1 that is incident on an array of static, equally-spaced particles when the restitution coefficient is nearly 1. After the first collision, the target particle moves with a slightly smaller speed, while the incident particle comes nearly to rest.

It is now helpful to exchange the identities of the two particles when they collide, so that the worldline of particle 1 is merely deflected slightly as it "passes through" particle 2. This same pattern continues in subsequent collisions so that the initial particle worldline gradually bends as it penetrates the array (Fig. 3.4). Let $v(n)$ be the velocity of the initial trajectory after n collisions. From the collision rule (3.6) and accounting for the change in particle labeling, $v(1) = 1 - \epsilon$. Similarly, to first order in ϵ, $v(n) = (1-\epsilon)^n \approx 1 - n\epsilon$. If n is sufficiently large, then the initial particle momentum is eventually exhausted and inelastic collapse has occurred. Setting $v(N) = 0$ then yields the critical collision parameter $\epsilon_c(N) \sim N^{-1}$ or critical restitution coefficient

$$r_c(N) \sim 1 - \frac{2}{N}\ . \tag{3.8}$$

Thus collapse occurs even for nearly elastic particles as N becomes sufficiently large — inelastic collapse becomes inevitable in the thermodynamic limit!

Maxwell Model for Free Cooling in One dimension

Because it is not feasible to solve the inelastic collision dynamics of a gas analytically, we turn to the simpler Maxwell model in which the collision rate is *independent* of particle velocities and positions — a particularly simple version of mean-field theory. Operationally, we pick two particles at random and define them to collide according to Eq. (3.6). Because of the simplicity of this collision dynamics, it is possible to solve the underlying Boltzmann transport equation and determine the velocity distribution analytically.

The Boltzmann transport equation for this Maxwell model model is

$$\frac{\partial P(v,t)}{\partial t} + P(v,t) = \int\int P(u_1,t)P(u_2,t)\delta\left[v - \epsilon u_1 - (1-\epsilon)u_2\right]\, du_1\, du_2\,. \tag{3.9}$$

This BTE conserves both the total number of particles, $\int P(v,t)\, dv = 1$, and the total momentum, $\int v\, P(v,t)\, dv = 0$. Because the time between every collision is the same, the elapsed time is proportional to the average number collisions that any particle experiences up to time t.

As an instructive starting point, we study low-order moments of the velocity distribution. Multiplying Eq. (3.9) by v^n, integrating over v, and performing some straightforward algebra, the rate equations for the

moments $M_n(t) = \int v^n\, P(v,t)\, dv$ are:

$$\dot{M}_n + a_n M_n = \sum_{m=2}^{n-2} \binom{n}{m} \epsilon^m (1-\epsilon)^{n-m} M_m M_{n-m} \tag{3.10}$$

for $n \geq 1$, with the coefficients $a_n(\epsilon) = 1 - \epsilon^n - (1-\epsilon)^n$. These equations may be solved recursively, starting with $M_0 = 1$ and $M_1 = 0$, and the first few non-trivial moments are:

$$\begin{aligned} M_2(t) &= M_2(0) e^{-a_2 t}, \\ M_3(t) &= M_3(0) e^{-a_3 t}, \\ M_4(t) &= \left[M_4(0) + 3M_2^2(0)\right] e^{-a_4 t} - 3M_2^2(t). \end{aligned} \tag{3.11}$$

The second moment quantifies velocity fluctuations through $M_2 \equiv T$, where this equivalence defines the effective temperature of the gas T. Since M_2 decays exponentially with time, the particles continuously slow down and eventually come to rest, with $P(v) \to \delta(v)$.

In stark contrast to equilibrium statistical mechanics, however, the temperature does not characterize the entire velocity distribution. The crucial point is that moments M_n with $n \leq 3$ obey scaling, but higher moments do not. For example, the leading asymptotic behavior of M_3 is $M_3 \sim e^{-a_3 t}$. Since $a_n(\epsilon) = 1 - \epsilon^n - (1-\epsilon)^n$, with $a_3 = 3a_2/2$, one can also write $M_3 \sim e^{-3a_2/2} \sim M_2^{3/2}$. Consequently, the second moment characterizes the scaling behavior of the third moment. However, $a_4 < 2a_2$ so that the ratio M_4/M_2^2 diverges as $t \to \infty$. In general, the moments scale as

$$M_n \sim \begin{cases} e^{-n a_2/2} & n \leq n^* \\ e^{-a_n t} & n \geq n^*, \end{cases} \tag{3.12}$$

with $n^* = 3$.

This change in scaling behavior is a sign that the velocity distribution has a power-law tail, rather than a Gaussian form. To determine the velocity distribution, we exploit the fact that the collision term in the Boltzmann equation is a convolution, so that the Fourier transform of this term is just a product. We thereby find that **why F? also give a derivation step** $F(k,t) = \int P(v,t)\, e^{ikv}\, dv$ evolves according to **following derivation is too rushed**

$$\frac{\partial F(k,t)}{\partial t} + F(k,t) = F((1-\epsilon)k, t)\, F(\epsilon k, t)\,. \tag{3.13}$$

We seek a scaling solution of the form $P(v,t) = T^{-1/2}\mathcal{P}(w)$, with the scaling variable $w = vT^{-1/2}$ to account for the velocity distribution approaching its final state in a self-similar fashion. The equivalent scaling of the Fourier transform is $F(k,t) = f(z)$ with $z = |k|\, T^{1/2}$, and the two scaling functions are related $f(z) = \int \mathcal{P}(w)\, e^{izw}\, dw$. Substituting the scaling ansatz into Eq. (3.13) and using the temperature decay rate $\frac{d}{dt}T = -\epsilon(1-\epsilon)T$, the scaling function $f(z)$ satisfies the ordinary differential equation

$$-\epsilon(1-\epsilon) z f'(z) + f(z) = f(z - \epsilon z) f(\epsilon z). \tag{3.14}$$

This equation is subject to the boundary conditions $f(z) \cong 1 - \frac{1}{2}z^2$, that follows from $F(k) \cong 1 - \frac{1}{2}k^2 T$.

The solution to (3.14) is **explain!!**

$$f(z) = (1+z)\, e^{-z}. \tag{3.15}$$

Inverting this Fourier transform, the scaled velocity distribution is[2]

$$\mathcal{P}(w) = \frac{2}{\pi}\, \frac{1}{(1+w^2)^2}. \tag{3.16}$$

Thus the velocity distribution of the inelastic Maxwell model gas has a power-law tail, $\mathcal{P}(w) \sim w^{-4}$ for $w \gg 1$, whose exponent is independent of the collision parameter ϵ.

The same general approach can be pursued in spatial dimension $d > 1$ but the details are much more complicated. The main point, however, is that the velocity distribution again has a power-law tail, but with the exponent now dependent on d and the collision parameter ϵ. the

[2]The inverse Fourier transform of $e^{-\kappa z}$ is $\frac{1}{\pi}\frac{\kappa}{\kappa^2+w^2}$; the inverse transforms of $z^n e^{-\kappa z}$ can be obtained using successive differentiation with respect to κ.

External Forcing

When energy is continuously injected into an inelastic gas, a steady state is reached as the energy input balances the energy loss in the inelastic collisions. The most natural way to realize such an energy input is by putting an inelastic gas in a closed container and shaking it. Experiments on this type of system indicate that the velocity distribution generically has a non-Maxwellian tail for a host of geometries and energy injection mechanisms. Here we discuss how to obtain this tail in the framework of the Maxwell model of random collisions.

We model the effect of the energy input by white-noise forcing in which each particle experiences an acceleration

$$\frac{dv_j}{dt} = \xi_j,$$

due to random noise amplitude that has zero mean, $\langle \xi_j \rangle = 0$, and no correlations in time, $\langle \xi_i(t)\xi_j(t')\rangle = D\delta_{ij}\delta\,(t-t')$. This white-noise forcing is equivalent to diffusion in velocity space with diffusion coefficient D. To account for the external forcing, the Boltzmann equation (3.9) should be augmented by a diffusion term in velocity space; that is

$$\frac{\partial P(v,t)}{\partial t} \to \frac{\partial P(v,t)}{\partial t} - D\frac{\partial^2 P(v,t)}{\partial v^2}\ .$$

With this additional term, the steady state velocity distribution $P(v) \equiv P(v, t=\infty)$ now satisfies:

$$\left(1 - D\,\frac{\partial^2}{\partial v^2}\right) P(v,t) = \int\int P(u_1)P(u_2)\delta\left[v - \epsilon u_1 + (1-\epsilon)u_2\right]\, du_1\, du_2. \tag{3.17}$$

The temperature changes according to $dT/dt + 2p(1-\epsilon)T = 2D$ **where does this come from?** so the steady state temperature is $T = D/[\epsilon(1-\epsilon)]$ and the relaxation toward the steady state is exponential, $|T - T_\infty| \sim \exp(-\text{const.} \times t)$. In the steady state, the Fourier transform $F(k) \equiv F(k, t=\infty)$ satisfies, following (3.13):

$$F(k) = (1+Dk^2)^{-1} F\left((1-\epsilon)k\right)\, F(\epsilon k). \tag{3.18}$$

To determine the large-velocity tail of the distribution, we solve (3.18) iteratively by repeatedly substituting the left-hand-side into the right-hand-side. Using the boundary conditions $F(0) = 1$ and $F'(0) = 0$ imposed by the conservation of the total particle number and the total momentum, the solution is

$$F(k) = \prod_{l=0}^{\infty}\prod_{m=0}^{l}\left[1 + \epsilon^{2m}(1-\epsilon)^{2(l-m)}Dk^2\right]^{-\binom{l}{m}}\ . \tag{3.19}$$

To extract the form of the high-velocity tail, we use the fact that the Fourier transform has an infinite series of poles located at $\pm i\left[p^{2m}(1-\epsilon)^{2(l-m)}\,D\right]^{-1/2}$. The simple poles at $\pm i/\sqrt{D}$ closest to the origin imply an exponential decay of the velocity distribution

$$P(v) \simeq \frac{A(\epsilon)}{v_*}\ e^{-|v|/v_*} \tag{3.20}$$

with $v_* = \sqrt{D}$ when $|v| \to \infty$ in which the dependence on the dissipation parameter appears only in the prefactor. In the quasi-elastic limit $\epsilon \to 1/2$, $A(\epsilon) \propto \exp[\pi^2/(12\epsilon)]$.

The leading behavior of the high-energy tail can be alternatively obtained using a useful and generic heuristic argument. For sufficiently large velocities, the gain term in the collision integral in Eq. (3.17) is negligible. The resulting equation for the steady state distribution

$$D\,\frac{d^2}{dv^2}P(v) = -P(v) \tag{3.21}$$

yields the exponential high-energy tail (3.20). This argument applies to arbitrary collision rates. For example, if $K(u_1,u_2) \propto |u_1 - u_2|^\delta$ for $|u_1 - u_2| \to \infty$, then the right-hand side in (3.21) becomes $-|v|^\delta P_\infty$ implying that $P_\infty(v) \propto \exp(-|v|^\gamma)$ with $\gamma = 1 + \delta/2$. For hard spheres Boltzmann equation ($\delta = 1$) one finds $\gamma = 3/2$.

Another important observation is that in the quasi-elastic limit, the velocity has a substantial Maxwellian core. In this limit, the Fourier transform (cumulant-1d) reads $F(k) \sim \exp[-Dk^2/\epsilon]$, so the velocity distribution is Gaussian. Comparing with the universal, dissipation-independent, tail behavior (3.20) shows that there is a cross-over velocity $v \sim \epsilon^{-1}$ marking the transition from the core to the tail. In summary,

$$P(v) \sim \begin{cases} \exp(-\epsilon v^2) & |v| \ll \epsilon^{-1}; \\ \exp(-|v|) & |v| \gg \epsilon^{-1}. \end{cases} \tag{3.22}$$

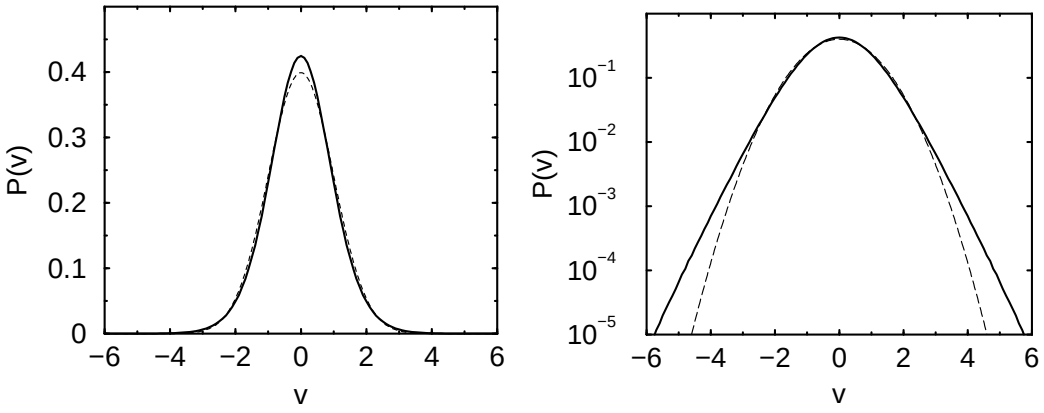


Figure 3.5: The steady state distribution versus the Maxwell-Boltzmann distribution. Shown is $P(v)$ versus $v \equiv v_x$ in the forced case for $d = 2$ and $r = 0$ (solid line). Also shown is the Maxwell-Boltzmann distribution (3.3) (dashed line). Both distributions are normalized such that $\langle v^2 \rangle = 1$. The distribution was obtained using a Monte-Carlo simulation of the inelastic collision process with white noise forcing.

3.4 Agglomeration

A simple and beautiful model is *ballistic agglomeration* in which particles move in straight-line trajectories and irreversibly stick together whenever a collision occurs. Such a model provides an idealized description of the large-scale agglomeration of matter in the universe to form planets and stars. We start with a gas of compact spherical objects that are randomly distributed in space, with masses m_i and radii proportional to $m_i^{1/3}$, and initial velocities $\mathbf{v}_i$. The initial values of m_i and v_i are drawn from some prescribed distributions. When the separation between two aggregates is less than the sum of their radii, they are defined to collide to form a larger aggregate, with the mass and the momentum conserved in the collision. That is

$$(m_1, \mathbf{p}_1) + (m_2, \mathbf{p}_2) \to (m_1 + m_2, \mathbf{p}_1 + \mathbf{p}_2). \tag{3.23}$$

The resulting aggregate is also assumed to maintain a spherical shape with a volume proportional to its mass.

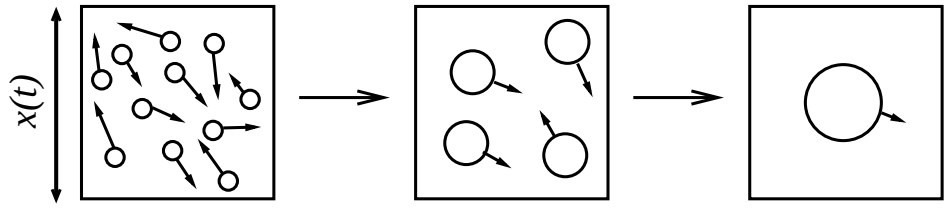


Figure 3.6: Illustration of ballistic agglomeration. The state of the system is illustrated at an early stage (left) and a late stage (right).

We anticipate that the aggregate growth proceeds in a self-similar manner, as illustrated in Fig. 3.6. There are several heuristic arguments one can formulate to determine the time dependence of the process. We give an argument that is based on elementary kinetic theory. A key assumption is that the momenta of aggregates remain uncorrelated throughout the process. Consequently the momentum of an aggregate of mass m is just the sum of m uncorrelated momenta. Thus $p \sim m \sim m^{1/2}$ and the aggregate velocity is $v \sim p/m \sim m^{-1/2}$. If there exists a typical aggregate mass, then after one collision time τ, two aggregates of approximately the same mass will meet, so that the typical mass should grow by of the order of itself. Thus the growth of aggregates is described by the rate equation

$$\frac{dm}{dt} \sim \frac{m}{\tau} \; . \tag{3.24}$$

We estimate τ from the elementary kinetic theory criterion that a collision occurs when the density times the volume of the collision tube swept out by the trajectory of an aggregate equals one: $n\ell R^{d-1} = 1$, where n is the aggregate density, $\ell \sim v\tau$ is the mean-free length, and R is the aggregate radius. Rewriting all these quantities in terms of the typical aggregate mass we have

$$\tau \sim \frac{1}{nvR^{d-1}} \sim m \times m^{1/2} \times m^{-(d-1)/d} \sim m^{(d+2)/2d} \; .$$

Finally, using this dependence for τ in the rate equation (3.24), we obtain

$$m \sim t^{2d/(d+2)}, \qquad v \sim t^{-d/(d+2)}, \qquad n \sim t^{-2d/(d+2)} \; . \tag{3.25}$$

Is the heuristic argument correct? Numerical simulations unambiguously support the result $\alpha = 2/3$ in one dimension. In this case, the theoretical argument stands on a firm ground: as agglomerates grow, they encompass a growing linear segment of the initial conditions. The corresponding region in space is indeed compact, so the constituent particles must be uncorrelated. In higher-dimensions the situation is less clear. The exponent $\alpha = 1$ certainly is close to the numerical results but recent simulations may be indicating a departure from this law. Such a departure may be due to the blobs acquiring a nontrivial non-compact geometry.

3.5 Jamming Traffic

An amusing application of kinetic theory is to traffic on rural highways, where often there is a single lane in each direction with no passing allowed. As many of us have experienced, faster cars accumulate behind slower-moving vehicles and significant clustering can arise. Each cluster is led by a slow vehicle that will ultimate catch up to a still-slower cluster if the road is sufficiently long. We can view traffic as a one-dimensional gas that evolves from a homogeneous to a clustered state because of the no-passing constraint.

We describe the traffic clustering phenomenon by the following idealized model. Each car has its own intrinsic speed at which it would move on an empty road. When a faster car catches up to a slower car, both cars subsequently move with speed the of the slower car (Fig. 3.7). Generally, when a cluster of m_2 cars, all moving at speed v_2, catches up to a cluster of m_1 cars moving at speed $v_1 < v_2$, the resulting "collision" leads to a cluster of $m_1 + m_2$ cars that moves at speed v_1. This clustering is described by the collision rule

$$(m_1, v_1) + (m_2, v_2) \to (m_1 + m_2, v_1). \tag{3.26}$$

For simplicity, we assume: (i) instantaneous collisions, (ii) sizeless cars, (iii) the initial car speeds that are drawn from an intrinsic distribution $P(v, t = 0)$, and (iv) initial car spacings are drawn from an independent exponential distribution with mean spacing equal to 1. Once the initial car speeds and positions are specified, traffic evolves deterministically according to the collision rule (3.26). The only source of randomness is the initial conditions. We are interested in the properties of the traffic, averaged over all initial configurations, or alternatively, in an infinite-size system.

Let's begin by giving a heuristic argument to obtain the basic characteristics of clustered traffic, namely, the time dependence of the typical cluster mass m, the typical separation between clusters ℓ, and the

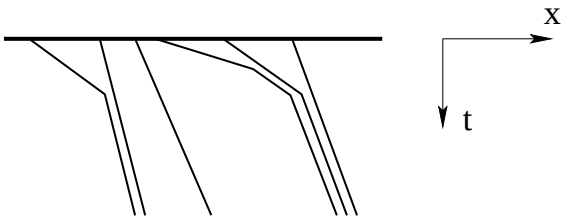


Figure 3.7: Traffic with no passing. Shown are the world lines of position x versus time t for initially unclustered traffic.

typical cluster speed v. As we shall see, these quantities are determined by the behavior of the initial speed distribution of the slowest cars. As a general example, let us assume the initial distribution of speeds[3]

$$P(v,0) \sim a\, v^{\mu} \tag{3.27}$$

as $v \to 0$.

On dimensional grounds, the quantities ℓ and v are related by $\ell \sim vt$. Because the separation between clusters should scale as the inverse of the concentration of clusters, which, in turn, is proportional to their inverse mass, we also have $\ell \sim m^{-1}$. Now let's relate the mass and speed of a cluster. Consider a car with speed v. The probability that a cluster of size k forms behind this car is given by $P_< P_>^k$, where $P_< = \int_0^v P(v',0)\, dv'$ and $P_> = 1 - P_<$ are the respective probabilities to find a car with speed less than or greater than v. Then the average size $\langle m(v)\rangle$ of the cluster behind a car of speed v is

$$\langle m(v)\rangle = \sum_{k=1}^{\infty} k\, P_< P_>^k = \frac{P_>}{P_<}\ .$$

For the power-law initial speed distribution $P(v,0) \sim av^{\mu}$ as $v \to 0$, we then find $\langle m(v)\rangle \sim v^{-1-\mu}$. Finally, we combine this relation with $m \sim vt$ to find

$$\begin{aligned} m &\sim t^{\alpha} \qquad & \alpha &= \frac{\mu+1}{\mu+2} \\ v &\sim t^{-\beta} \qquad & \beta &= \frac{1}{\mu+2}. \end{aligned} \tag{3.28}$$

The decay exponents satisfy the scaling relation $\alpha + \beta = 1$ which merely reflects the dimensions of the basic relations $\ell \sim vt$ and $\ell \sim m^{-1}$.

Now let's turn to the speed distribution of clusters. For simplicity, consider first the special case where cars have only 2 possible speeds, v_1 and v_2, with respective probabilities $P_1(0)$ and $P_2(0)$. Slow cars never catch any other vehicle, so their density is conserved, $P_1(t) = P_1(0)$. Fast cars move at their intrinsic speed v_2 before colliding a slower car. To avoid a collision up to time t, a fast car must have a segment of length $(v_2 - v_1)t$ ahead of its initial position that is free of slow cars. Since the initial spatial distribution of cars is random, this exclusion probability decays exponentially with the interval length $\exp[-P_1(0)(v_2 - v_1)t)]$. Therefore, the density of the faster cars is

$$P_2(t) = P_2(0) \exp\left[-P_1(0)(v_2 - v_1)t\right]. \tag{3.29}$$

Now consider traffic with cars that move at three distinct intrinsic speeds $v_1 < v_2 < v_3$. Clusters with speeds v_1 and v_2 are simply unaffected by the presence of faster cars so the previous conclusions for their densities $P_1(t)$ and $P_2(t)$ hold! On the other hand, for a fast car to maintain a speed v_3, it must avoid colliding with both clusters of speeds v_1 *and* v_2. The probability for these two independent events is given by a product of the two "exclusion probabilities"

$$P_3(t) = P_3(0) \exp[-P_1(0)(v_3 - v_1)t] \exp[-P_2(0)(v_3 - v_2)t]. \tag{3.30}$$

[3]Without loss of generality, we subtract the speed of the slowest car from all speeds so that the minimum speed is 0.

We may now generalize the cluster speed distribution for an arbitrary initial speed distribution:

$$P_n(t) = P_n(0) \exp\left[-t \sum_{i=1}^{n-1} P_i(0)(v_n - v_i)\right], \tag{3.31}$$

where $P_n(0)$ is the probability that a car has an initial speed v_n for a discrete set of speeds $\{v_n\}$, with $v_1 < v_2 < v_3 < \ldots$. In the continuum limit, Eq. (3.31) now becomes

$$P(v,t) = P(v,0) \exp\left[-t \int_0^v dv'\,(v - v')P(v',0)\right]. \tag{3.32}$$

Thus the density of cars of any positive speed decays exponentially in time, with a decay rate that is a growing function of v — the faster the intrinsic speed of a car, the more likely it will become stuck behind a bus.

We can express Eq. (3.32) in a Boltzmann-like form by differentiating the logarithm of this equation with respect to time to give the *linear* evolution equation

$$\frac{\partial P(v,t)}{\partial t} = -P(v,t) \int_0^v dv'(v - v')P(v',0). \tag{3.33}$$

As in the classic Boltzmann equation (3.4), there is an integration over all possible collision partners with speeds $v' < v$ in which the collision rate is proportional to the relative speed $|v - v'|$. There does not exist, however, a gain term because a cluster of a given speed cannot be created if it doesn't already exist. However, traffic theory has a fundamental difference with classical kinetic theory. The evolution equations of traffic, Eqs. (3.32) of (3.33) are *non-local* in time because of the perpetual memory of the initial conditions: the speed distribution at time t is expressed in terms of the initial speed distribution. **paragraph incomplete**

Substituting the initial distribution of speeds $P(v,0) \sim a\, v^\mu$ as $v \to 0$. in Eq. (3.32), we find, in the long-time limit,

$$P(v,t) \sim a\, v^\mu \, \exp(-btv^{\mu+2}), \tag{3.34}$$

with $b = a/[(\mu+1)(\mu+2)]$. Notice that the speed distribution can be written in the scaling form

$$P(v,t) \simeq t^{\beta-\alpha}\Phi(vt^\beta). \tag{3.35}$$

However, the speed distribution near $v = 0$ is an invariant of the dynamics. No matter how long the time, a small number of cars have yet to encounter still slower cars and $P(v,t) \sim P_0(v)$ for these slow cars. From this speed distribution, the leading asymptotic behaviors of the concentration, $c(t) = \int P(v,t)\,dv$, and the average speed, $\langle v(t)\rangle = c^{-1} \int v\, P(v,t)\,dv$, are simply

$$\begin{aligned} c(t) &\sim A\, t^{-\alpha} \qquad \text{with} \qquad A = (\mu+1)b^\beta \Gamma(\alpha), \quad \alpha = \frac{\mu+1}{\mu+2} \\ \langle v(t)\rangle &\sim B\, t^{-\beta} \qquad \text{with} \qquad B = \frac{b^{-\beta}}{\Gamma(\alpha)}, \qquad\qquad \beta = \frac{1}{\mu+2}. \end{aligned} \tag{3.36}$$

The asymptotic behavior reflects the initial conditions very strongly. The exponents are non-universal and are dictated by a particular aspect of the initial conditions, namely, the form of the distribution of the slowest cars. This of course reflects the nature of the collision rule as the slowest cars (unfortunately) govern the congestion.

I don't know how to derive the correlation function. Moreover, the joint probability density $P_2(v,v';t)$, the probability of finding two particles of speed v and speed v' at time t, can be read immediately from the integrand

$$P(v,v';t) = P(v_>,t)P(v_<,0) \tag{3.37}$$

with $v_> = \max(v,v')$ and $v_< = \min(v,v')$. This joint density differs from the traditional mean-field theory (the *stossansatz*) where $P(v,v';t) = P(v,t)P(v',t)$. Such factorization does not hold because spatial correlations build-up dynamically. Even though the initial state contains no correlations between the speeds of the particles and their positions, such correlations do eventually develop. Intuitively, space-speed correlations may reflect regions where particles have very close speeds and thus experience very few collisions due to this "shielding" effect.

Problems

1. Verify the general behavior of the moments quote in Eq. (3.12) for the one-dimensional inelastic gas by exploiting the inequality $a_n < a_m + a_{n-m}$ for all $1 < m < n-1$.inelastic.
2. Large moment as a characteristic of the tails of the distribution. Show that the for the 1D freely cooling inelastic Maxwell model, kinetics of large moments $\lim_{n\to\infty} M_n(t)$ coincide with the behavior of the tail of the distribution $\lim_{v\to\infty} P(v,t)$.
3. The quasi-elastic limit as a singular perturbation. Expand the rate equation (3.13) to first order in ϵ and then solve it. Show that the emerging solution remembers the initial conditions forever.
4. Development of singularities in compact velocity distributions. Express the Fourier transform $F(k,t)$ explicitly in terms of the initial distribution $F_0(k)$. Introducing the transformations $G(k,t) = e^t F(k,t)$ and $\tau = 1 - e^{-t}$, reduce the evolution equation (3.13) to $G_\tau(k) = G(\epsilon k)G(k-\epsilon k)$ and express then, express the solution as a formal Taylor series in powers of τ^n.
5. Estimating the large-velocity tail σ. Find the leading asymptotic behavior of the exponent σ as $d \to \infty$. To perform this asymptotic analysis, note that for large dimensions, the integral $\langle \eta^{(\sigma-d)/2} \rangle$ vanishes exponentially with the dimension d and then write $\sigma/d = f(\epsilon)$.
6. Obtain the solution (cumulant-1d) from the solution (3.19).

Solutions

1. The large moments decay universally as $M_n \sim \exp(-t)$ because $a_n \to 1$ as $n \to \infty$. For the tail of the velocity distribution, the gain term is negligible in the rate equation (3.9) (Intuitively, large velocities can only shrink due to collisions). Thus, $\frac{d}{dt}P(v,t) = -P(v,t)$ and the large-v tail of the velocity distribution also decays as $P(v,t) \sim P(v,0)\exp(-t)$.
2. The equation is $F_t + pkF_k = 0$ and using the method of characteristics, the solution is $F(k,t) = F_0(ke^{-pt})$. It remembers the initial conditions forever, in contradiction with the similarity solution (3.16).
3. The formal Fourier expansion reads $F(k,t) = e^{-t}\sum_{n=0}^{\infty} \frac{(1-e^{-t})^n}{n!} F_n(k)$. The expansion functions $F_n(k)$ are obtained from the recursion relation $F_{n+1}(k) = \sum_{m=0}^{n} \binom{n}{m} F_m(k-\epsilon k)F_{n-m}(\epsilon k)$ and $F_0(k) \equiv F(k, t=0)$. The expansion functions are products of F_0 with stretched arguments of the form $k\epsilon^l(1-\epsilon)^m$. This implies that starting from a compact initial distribution $P(v,0)$, the velocity distribution $P(v,t)$ develops a set of singularities. For instance, a distribution with support in $[-v_0, v_0]$ becomes non-analytic at an infinite set of points $v_{l,m} = \pm\epsilon^l(1-\epsilon)^m v_0$.
4. $f(\epsilon) = \frac{1+\frac{3}{2}\epsilon-\epsilon^3-\epsilon^{1/2}\left(1+\frac{5}{4}\epsilon\right)^{1/2}}{\epsilon(1-\epsilon^2)}$.
5. Assume that the restitution coefficient is randomly chosen according to the distribution $f(\epsilon)$. Generalize Eq. (refroot).

1. Heuristic derivation of extremal statistics. Evaluate the limiting behaviors of the scaling function $\Phi(z)$ in Eq. (3.35) when $z \to 0$ and $z \to \infty$. This scaling form is consistent with the exact solution (3.32) with the scaling function (problem 2)

$$\Phi(z) = az^\mu \exp\left[-bz^\mu\right]. \tag{3.38}$$

2. Evaluate the cluster size distribution $P(v,t)$ and the cumulative distribution $Q_m(v,t)$ for the special case $P_0(v) = e^{-v}$. Show that the result is consistent with the general scaling behavior.
3. Uniform final distribution. Obtain the initial speed distribution $P_0(v)$ and the flux for the final speed distribution $P(v) = c$ for $0 < v < 1$. What is the corresponding μ?
4. What is the scaling distribution in the ballistic aggregation model. Evaluate the tail of the speed distribution using the Lifshitz tail argument.

Solutions

1. The limit $z \to 0$ is discussed in the text, $P(v,t) \simeq P_0(v)$. The limit $z \to \infty$ is obtained from the Lifshitz tail argument (see matters of technique and freely cooling inelastic gases) by noting that $P(v,t) \sim \exp(-Ct)$ for fast particles. An interval of the length t has to be empty ahead in the initial configuration.

2. The cluster size distribution is $P(v,t) = \exp[-v - t(e^{-v} + v - 1)]$. The cumulative distribution is

$$Q_m(v,t) = P(v,t)e^{-(m-1)v}t^{n-1}\frac{\Gamma(t+1)}{\Gamma(t+m)}.$$

   These solutions are consistent with the scaling behavior for the case $a = 1$ and $\mu = 1$.

3. The initial speed distribution is $P_0(v) = R^{-1} + \frac{1}{2}cv^2$. The flux is $J = [(3+\lambda)\sqrt{\lambda}\tan^{-1}\sqrt{\lambda} + \lambda - \ln(1+\lambda)]/R$ with $\lambda = \frac{1}{2}Rc = \frac{3}{2}[\sqrt{1+2R/2} - 1]$. The corresponding $\mu = 2$ and $J \sim R^{-1/4}$.

4. The speed distribution has the scaling form $P(v,t) \simeq t^{-\alpha+d\beta}\Phi(vt^{\beta})$. The tail of the speed distributing is $\Phi(z) \sim \exp(-z^{\gamma})$ with $\gamma = \frac{d+2}{d}$.

Chapter 4

AGGREGATION

In aggregation, reactive clusters join irreversibly whenever two of them meet. Aggregation is ubiquitous in nature: it underlies milk curdling, blood coagulation, and star formation by gravitational accretion. Aggregation also provides a beautiful example of many paradigmatic features of non-equilibrium phenomena, such as scaling, phase transitions, and non-trivial steady states. Schematically, we write aggregation as

$$A_i + A_j \xrightarrow{K_{ij}} A_{i+j},$$

in which a cluster of mass $i+j$ is created at an intrinsic rate K_{ij} by the aggregation of two clusters of mass i and mass j. The goal of this chapter is to determine the concentration of clusters of mass k at time t, $c_k(t)$, and to understand which features of the underlying reaction rate, or kernel, K_{ij} influence this distribution.

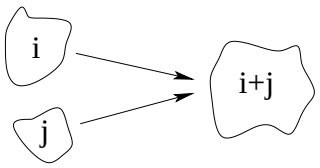


Figure 4.1: Clusters of mass i and mass j merge irreversibly into a cluster of mass $i+j$.

The Master Equations

The starting point for treating aggregation is the infinite set of master equations that describes how the cluster mass distribution changes with time. These master equations are

$$\frac{dc_k(t)}{dt} = \frac{1}{2}\sum_{\substack{i,j\\ i+j=k}} K_{ij}\, c_i(t)\, c_j(t) - c_k(t)\sum_{i=1}^{\infty} K_{ik}\, c_i(t). \tag{4.1}$$

The first term on the right-hand side of (4.1) describes the gain in the concentration of clusters of mass $k = i + j$ due to the coalescence of clusters of mass i and mass j. The second term accounts for the loss of clusters of mass k due to their reaction with other clusters. In the approximation of well-mixed reactants, the rate at which an i-mer and j-mer meet is $K_{ij}c_ic_j$, and the prefactor $\frac{1}{2}$ in the gain term ensures the correct accounting of reactions between same-mass clusters.[1]

In equations (4.1), and generally throughout chapter we tacitly assume that the mass k runs over the integers — this merely implies that we measure mass in terms of a minimal mass, and a cluster of mass k

[1]It is helpful to consider a finite system to understand this factor. Denote by N_k the total number of clusters of mass k. For $i \neq j$ there are N_iN_j pairs of type ij, while the number of same-mass pairs is $\frac{1}{2}N_k(N_k-1) \to \frac{1}{2}N_k^2$ in the thermodynamic limit. Thus the prefactor $\frac{1}{2}$ properly accounts for the the relative fraction of same-mass pairs. The loss term for same-mass pairs in (4.1) is $K_{kk}c_kc_k$ rather than $\frac{1}{2}K_{kk}c_kc_k$ since *two* clusters of mass k disappear in such a collision.

contains k primal, minimal-mass clusters. Primal clusters are called *monomers* while clusters of mass k are termed $k-$mers. With this convention regarding the mass, the reaction rates form an infinite symmetric matrix $K_{ij} = K_{ji}$. The master equations (4.1) admit an important integral of motion — the mass density

$$M(t) = \sum_{k\geq 1} k\, c_k(t) \tag{4.2}$$

is conserved. To verify this conservation law we write

$$\frac{dM}{dt} = \sum_k k\,\frac{dc_k}{dt} = \sum_k \sum_{i+j=k} \frac{1}{2} K_{ij}\,(i+j)\,c_i\,c_j - \sum_k \sum_i K_{ik}\,k\,c_i\,c_k = 0. \tag{4.3}$$

The outer sum over k causes the sums over i and j in the gain term to become independent and unrestricted. The gain and loss terms then cancel and therefore the mass density is manifestly conserved.

The master equations are the starting point in almost all studies of aggregation, and it is instructive to highlight the assumptions underlying this approach, including:

- The system is well mixed, and the reaction proceeds with a rate proportional to the product of reactant densities. This is the *mean-field* assumption.
- Bimolecular reactions. The system is sufficiently dilute so that higher-body interactions are negligible.
- Shape independence. The aggregate mass is the only dynamical variable; cluster shape play no role in the evolution. One such example is the aggregation of spherical liquid droplets.
- Thermodynamic limit. The system is sufficiently large that cluster concentrations are continuous functions; discreteness effects are ignored.

4.1 Exact Solutions

The master equations are a formidable infinite set of coupled non-linear differential equations that are soluble only for a few neat kernels. Many clever solution techniques have been developed for these kernels and we present several such approaches. We start with the constant reaction kernel because it represents an ideal playground to illustrate a variety of approaches. We then turn to more challenging cases of the product and sum kernels, $K_{ij} = ij$ and $K_{ij} = i + j$, respectively. These three examples represent most of the exactly solved models of aggregation.

Constant Reaction Rates

The constant kernel aggregation was proposed and solved in the first paper about aggregation (Smoluchowski, 1917). A crude physical justification of the model is based on the form of the reaction kernel for Brownian aggregation. From Sec. 2.5, the reaction rate for spherical aggregates that undergo Brownian motion is [see Eq. (2.51)]

$$\begin{aligned} K_{ij} &\sim (D_i + D_j)(R_i + R_j) \propto (i^{-1/3} + j^{-1/3})(i^{1/3} + j^{1/3}) \\ &= 2 + \left(\tfrac{i}{j}\right)^{1/3} + \left(\tfrac{j}{i}\right)^{1/3}. \end{aligned} \tag{4.4}$$

The Brownian kernel — as yet unsolved — shares one important feature with the constant kernel — they both are invariant under the transformation $(i, j) \longrightarrow (ai, aj)$, that is, $K_{ai,aj} = K_{i,j}$. This suggests that the constant kernel is a reasonable but uncontrolled approximation for the physically-important Brownian kernel.

For the constant kernel, we choose $K_{ij} = 2$ for convenience, and then the master equations are

$$\frac{dc_k}{dt} = \sum_{i+j=k} c_i c_j - 2c_k \sum_{i=1}^{\infty} c_i \equiv \sum_{i+j=k} c_i c_j - 2c_k\, N \tag{4.5}$$

where $N(t) = \sum_{k\geq 1} c_k(t)$ is the concentration of clusters of any mass. The first few of these equations are

$$\begin{aligned}
\dot{c}_1 &= -2c_1\, N \\
\dot{c}_2 &= c_1^2 - 2c_2\, N \\
\dot{c}_3 &= 2c_1\, c_2 - 2c_3\, N \\
\dot{c}_4 &= 2c_1\, c_3 + c_2^2 - 2c_4\, N \\
\dot{c}_5 &= 2c_1\, c_4 + 2c_2\, c_3 - 2c_5\, N \\
\dot{c}_6 &= 2c_1\, c_5 + 2c_2\, c_4 + c_3^2 - 2c_6\, N\,,
\end{aligned} \tag{4.6}$$

where the overdot denotes the time derivative.

One major lesson that emerges from studies of aggregation and other irreversible processes is that the asymptotic behavior (which is the most interesting characteristic of the system) depends on the initial condition in a trivial way, *e.g.*, in terms of the entire mass, while the detailed behavior of the initial data is irrelevant. Therefore it is convenient to choose the simplest initial condition to avoid cluttered formulae. In the context of aggregation, the monomer-only initial condition

$$c_k(0) = \delta_{k,0} \tag{4.7}$$

is the most natural and simplest choice. If not stated otherwise, we shall always assume such an initial condition in the following. Before solving the initial-value problem (4.5)–(4.7), let us look at the moments of the mass distribution, where much information can be gleaned with relative little effort.

Moments

For master equations with neat kernels, the moments $M_n(t) \equiv \sum_{k\geq 1} k^n c_k(t)$ usually satisfy simple rate equations that may be solvable even if the master equations are unsolvable. Moments also immediately give us some basic information about the mass distribution, *e.g.*, the ratio $M_1/M_0 \equiv M/N$ gives an estimate for the average cluster mass.

In the case of the constant reaction rates, the moment equations are particularly simple. Using Eqs. (4.5) we deduce

$$\begin{aligned}
\frac{dM_n}{dt} = \sum_{k=1}^{\infty} k^n\, \dot{c}_k &= \sum_{k=1}^{\infty} k^n \Big[\sum_{i+j=k} c_i c_j - 2c_k \sum_{i=1}^{\infty} c_i \Big] \\
&= \sum_{i,j}^{\infty} (i+j)^n\, c_i\, c_j - 2M_n\, M_0,
\end{aligned} \tag{4.8}$$

where the sums over i and j are unrestricted in the second line. The explicit equations for the first few moments are

$$\begin{aligned}
\dot{M}_0 &= \sum_{i,j} c_i\, c_j - 2M_0^2 = -M_0^2 \\
\dot{M}_1 &= \sum_{i,j} (i+j)\, c_i\, c_j - 2M_1 M_0 = 0 \\
\dot{M}_2 &= \sum_{i,j} (i^2 + 2ij + j^2)\, c_i\, c_j - 2M_2 M_0 = 2M_1^2 \\
\dot{M}_3 &= \sum_{i,j} (i^3 + 3i^2 j + 3ij^2 + j^3)\, c_i\, c_j - 2M_3 M_0 = 6M_1 M_2 \\
\dot{M}_4 &= \sum_{i,j} (i^4 + 4i^3 j + 6i^2 j^2 + 4ij^3 + j^4)\, c_i\, c_j - 2M_4 M_0 = 8M_1 M_3 + 6M_2^2
\end{aligned} \tag{4.9}$$

For the monomer-only initial condition, $M_n(0) = 1$ for all $n \geq 0$. The solution for the zeroth moment $M_0 = N$ is

$$N(t) = \frac{1}{1+t}\,. \tag{4.10}$$

Solving equations (4.9) for the higher moments one by one we obtain $M_1 = 1$, $M_2 = 1+2t$, $M_3 = 1+6t+6t^2$, $M_4 = 1 + 14t + 36t^2 + 24t^3$, *etc.* In general, $M_n \simeq n!\, t^{n-1}$ as $t \to \infty$.

Pedestrian approach

The master equations (4.5) are recursive and therefore they can be solved one by one. For the monomer-only initial condition, we substitute $N(t)$ from (4.10) into the first of (4.6) and integrate to give $c_1(t) = (1+t)^{-2}$. Having found c_1, the master equation for c_2 becomes

$$\dot{c}_2 = (1+t)^{-4} - 2(1+t)^{-1}\, c_2$$

Solving this equation subject to $c_2(0) = 0$ gives $c_2(t) = t/(1+t)^3$. The next density satisfies

$$\dot{c}_3 = 2t(1+t)^{-5} - 2(1+t)^{-1}\, c_3\,, \qquad c_3(0) = 0$$

whose solution is $c_3(t) = t^2/(1+t)^4$. Continuing this recursive approach we find $c_4(t) = t^3/(1+t)^5$, then $c_5(t) = t^4/(1+t)^6$, *etc.* This pattern suggests the general solution

$$c_k(t) = \frac{t^{k-1}}{(1+t)^{k+1}}\,. \tag{4.11}$$

A direct argument using induction proves that this guess is correct.

Elimination of loss terms

One useful trick that often simplifies master equations is based on eliminating loss terms. For example consider the concentration ratio $\phi_k \equiv c_k/c_1$, whose master equation is readily shown to be

$$\frac{d\phi_k}{dt} = c_1 \sum_{i+j=k} \phi_k \phi_j. \tag{4.12}$$

Thus the loss term has indeed disappeared. We now define the rescaled time

$$\tau = \int_0^t dt'\, c_1(t') \tag{4.13}$$

so that the master equation reduces to $\phi_k' = \sum_{i+j=k} \phi_i\, \phi_j$, where the prime denotes differentiation with respect to τ. Solving for the first few ϕ_k, it is immediately clear that the solution is $\phi_k = \tau^{k-1}$. To relate τ and the time t, we substitute the already-established monomer density $c_1 = (1+t)^{-2}$ into (4.13) and find $\tau = t/(1+t)$. Finally, substituting into $c_k = \phi_k\, c_1$, we re-derive (4.11).

The elegant closed-form solution (4.11) has many interesting asymptotic properties, including

1. For $t \to \infty$, $c_k \to t^{-2}\, e^{-k/t}$. Thus for fixed k, each $c_k(t)$ approaches a common limit that decays as t^{-2} as $t \to \infty$ (Fig. 4.2). For $k < t$, the mass distribution is nearly flat, as shown on the right side of the figure.

2. The area under the mass distribution is therefore proportional to $t^{-2} \times t = t^{-1}$, which reproduces the correct time dependence of the total concentration of clusters.

3. The short- and long-time limits of c_k can be easily determined without solving the full master equations. For the short-time behavior we ignore the loss terms in the master equations. The resulting equations have the same form as Eqs. (4.12), from which we obtain $c_k(t) \sim t^{k-1}$ for $t \ll 1$. Conversely for $t \to \infty$, there is no production of k-mers for fixed k. We therefore ignore the gain terms in the master equation to give $\dot{c}_k \sim -2c_k N$, whose solution is $c_k \sim t^{-2}$.

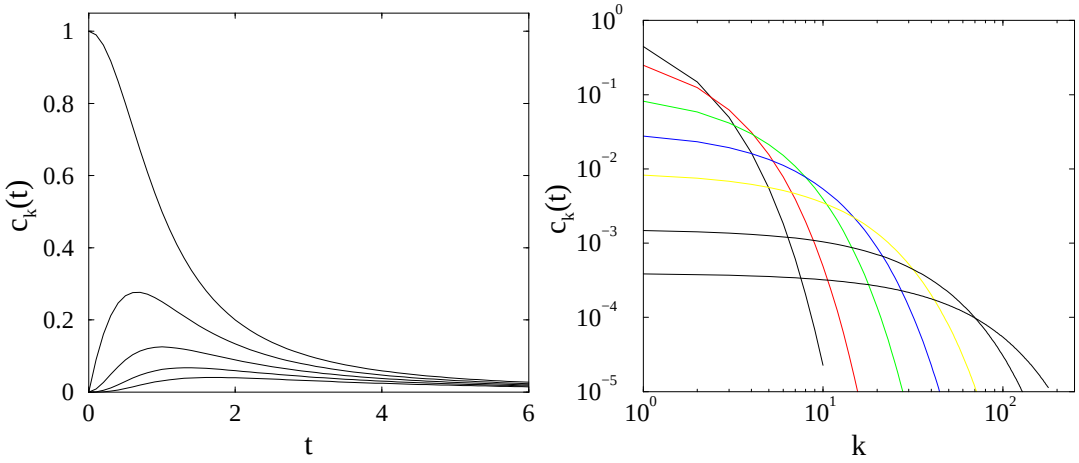


Figure 4.2: Left: Cluster concentrations $c_k(t)$ versus time for constant kernel aggregation for $k = 1, 2, 3, 4, 5$ (top to bottom). The concentrations approach a common limit as $t \to \infty$, as predicted by the scaling form in Eq. (4.11). Right: $c_k(t)$ versus k on a double logarithmic scale for $t = 1, 2, 5, 10, 20, 50$, and 100 (upper left to lower right).

Exponential ansatz

Solutions to the master equations often have an exponential form — equation (4.11) is one such example. By making use of this assumption at the outset, we can simplify the rate equations considerably. For the case of the constant kernel, the appropriate exponential ansatz is

$$c_k(t) = A(t)\, a(t)^{k-1}, \tag{4.14}$$

with the initial conditions $A(0) = 1$ and $a(0) = 0$. Choosing the power $k-1$ for a makes the ansatz compatible with the monomer-only initial condition. Substituting the ansatz (4.14) into the master equations (4.5), and dividing both sides of the equation by c_k, we find

$$\frac{\dot{A}}{A} + (k-1)\frac{\dot{a}}{a} = (k-1)\frac{A}{a} - \frac{2A}{1-a}.$$

Thus the exponential ansatz leads to k-dependent and k-independent components that we can equate separately to give

$$\dot{A} = -\frac{2A^2}{1-a}; \qquad \dot{a} = A. \tag{4.15}$$

If we had chosen a different power of a in the initial ansatz, there would not be the natural alignment of terms given above, but it would also be clear from the degree of misalignment how to choose the correct power of a. Since $\sum_{k\geq 1} kc_k = A\sum_{k\geq 1} ka^{k-1} = A(1-a)^{-2}$, mass conservation implies $A = (1-a)^2$; the same conservation law also follows from equations (4.15). Substituting $A = (1-a)^2$ back into (4.15) we immediately find

$$A = \frac{1}{(1+t)^2}; \qquad a = \frac{t}{1+t}, \tag{4.16}$$

thus reproducing the solution for $c_k(t)$ in Eq. (4.11).

The exponential ansatz has an advantage over the two previous approaches in that it involves less guesswork and it requires dealing with two (instead of infinitely many) differential equations. In addition, this ansatz works for all exponentially decaying initial conditions.

Example 1. *Exponentially decaying initial data.* Let $c_k(0) = (1-q)^2 q^{k-1}$ where $0 < q < 1$; the monomer-only initial condition is recovered for $q \to 0$. We use the exponential ansatz (4.14) with $A = (1-a)^2$ and now we must solve $\dot{a} = (1-a)^2$ subject to $a(0) = q$. We find

$$a(t) = 1 - \frac{1-q}{1+(1-q)t}, \qquad A(t) = \left[\frac{1-q}{1+(1-q)t}\right]^2$$

The parameter q does not affect the qualitative asymptotic behavior. For instance, for fixed mass, the densities approach the common limit t^{-2} as $t \to \infty$. When both mass and time diverge in such a way that the ratio k/t remains finite, the mass distribution attains a scaling form $c_k \simeq t^{-2}\exp(-k/t)$.

Generating function method

A powerful approach for solving the master equations is the generating function method. This technique is ideally-suited for aggregation because the master equations have a discrete convolution form that transform into an easily-soluble product by the generating function. The generating function is defined as

$$\mathcal{C}(z,t) \equiv \sum_{k=1}^{\infty} c_k(t) z^k, \tag{4.17}$$

and it encodes the entire mass distribution within a single function. To apply the generating function method to constant-kernel aggregation, we take each of the equations for c_k in (4.6), multiply by z^k, and sum over all k. This gives

$$\frac{d\mathcal{C}}{dt} = \sum_k \sum_{i+j=k} c_i z^i \; c_j z^j - 2\sum_k c_k z^k \; \sum_i c_i = \mathcal{C}^2 - 2\,\mathcal{C}N \tag{4.18}$$

Here we use the fact that the sum over k renders the two sums over i and j independent, so that the first term reduces to a product. This reduction to a product is the crucial simplification of the generating function. Since the rate equation for N is $\dot{N} = -N^2$, the function $\mathcal{C}_- \equiv \mathcal{C} - N$ satisfies $\dot{\mathcal{C}}_- = \mathcal{C}_-^2$. This equation should be supplemented with an initial condition which is $\mathcal{C}_-(z, t=0) = z - 1$, for the monomer-only initial condition. The solution is $\mathcal{C}_- = (z-1)/[1-(z-1)t]$, from which we obtain

$$\mathcal{C} = \frac{1}{1+t}\,\frac{z}{1-(z-1)t}\,. \tag{4.19}$$

Expanding (4.19) as a power series in z gives

$$\mathcal{C}(z,t) = \sum_{k=1}^{\infty} z^k \frac{t^{k-1}}{(1+t)^{k+1}}\,.$$

From this form, we directly read off the mass distribution and thereby recover Eq. (4.11).

For an *arbitrary* initial condition the generating function is

$$\mathcal{C}(z,t) = (1+t)^{-2}\,\frac{\mathcal{C}_0(z)}{1-\frac{t}{1+t}\mathcal{C}_0(z)}\,, \tag{4.20}$$

where $\mathcal{C}_0(z) = \mathcal{C}(z, t=0)$ and we also assume that $N(t=0) = \mathcal{C}_0(z=1) = 1$. Expanding the generating function (4.20) as a power series in z to obtain the densities $c_k(t)$ for all k is straightforward in principle but may be computationally tedious.

Example 2. *Initial data with finite support.* The monomer-only initial condition is the simplest example of initial data with finite support. The next simplest possibility is an initial state that consists of a mixture of monomers and dimers, that is $c_1(0), c_2(0) > 0$, with $c_1(0) + c_2(0) = 1$, while $c_j(0) = 0$ for $j \geq 3$. Then $\mathcal{C}_0(z) = c_1(0)z + c_2(0)z^2$ so that (4.20) gives

$$\mathcal{C}(z,t) = (1+t)^{-2}\,\frac{c_1(0)z + c_2(0)z^2}{1-\frac{t}{1+t}\left[c_1(0)z + c_2(0)z^2\right]}$$

To expand of this generating function in powers of z, we rewrite the quadratic polynomial in the denominator as $(1-z/z_1)(1-z/z_2)$, where $z_{1,2}(t)$ are the roots of the polynomial. Then we present $[(1-z/z_1)(1-z/z_2)]^{-1}$ as a combination of $(1-z/z_1)^{-1}$ and $(1-z/z_2)^{-1}$, and expand each of these factors as a geometric series. Explicit results can be similarly obtained if aggregation begins with a mixture of monomers, dimers, trimers, and 4−mers, while $c_j(0) = 0$ for $j \geq 5$. The denominator of the generating function is now a polynomial of degree 4 whose roots can be explicitly computed. Then writing $[(1-z/z_1)(1-z/z_2)(1-z/z_3)(1-z/z_4)]^{-1}$ as a combination of $(1-z/z_i)^{-1}$, explicit formulae could be obtained. Since a generic polynomial of degree 5 and higher cannot be factored, there do not exist explicit results for the mass distribution if the initial condition contains clusters of mass ≥ 5.

Despite the lack of explicit results for arbitrary initial mass distributions with finite support, the generating function solution (4.20) allows us to deduce the important *asymptotic* behaviors. Consider an initial mass distribution that vanishes for masses larger than m, that is, $c_j(0) = 0$ for $j \geq m+1$. While we cannot compute the roots of the polynomial

$$1 - \frac{t}{1+t}\,[c_1(0)z + c_2(0)z^2 + \ldots + c_m(0)z^m] = 0$$

for $m \geq 5$, we only need the the smallest root z_1 because it contribution, $(z_1)^{-k}$, dominates those of all other roots for large k. Because this smallest root is close to 1 in the long-time limit, we write $z_1 = 1+\epsilon$ and note that

$$\mathcal{C}_0(z) = c_1(0)z_1 + c_2(0)z_1^2 + \ldots + c_m(0)z_1^m = 1 + M(0)\epsilon + O(\epsilon^2),$$

where $c_1(0) + c_2(0) + \ldots + c_m(0) = 1$ is the initial cluster density and $c_1(0) + 2c_2(0) + \ldots + mc_m(0) = M(0)$ is the initial mass density. Using this result for $\mathcal{C}_0(z)$ in Eq. (4.20), the generating function becomes

$$\mathcal{C}(z,t) \approx \frac{1}{1+t}\;\frac{1+M(0)\epsilon}{1-tM(0)\epsilon} \simeq \frac{1}{t^2M(0)z_1}\;\frac{1}{1-z/z_1},$$

where the smallest root of the generating function $z_1 = 1 + 1/[M(0)t]$. Now expanding in a power series in z, we obtain the scaling form for the mass distribution

$$c_k \simeq \frac{1}{M(0)\,t^2}\;e^{-k/[M(0)t]}\;. \tag{4.21}$$

Example 3. *Algebraically decaying initial data*. If the initial mass distribution is unbounded and decays slowly, pathological behavior may occur. One such example is

$$c_k(0) \to \frac{C}{k^\alpha} \qquad \text{when} \quad k \gg 1. \tag{4.22}$$

The exponent α should be larger than one, since the zeroth moment of the mass distribution must converge — otherwise the master equations (4.5) are ill-defined. Apart from this restriction, the decay exponent α is arbitrary. In particular, if $\alpha \leq 2$, the first moment diverges, $M(0) = \infty$, and the asymptotic (4.21) no longer holds.

Let $1 < \alpha < 2$ (the marginal $\alpha = 2$ case is more tedious due to the presence of logarithms). The generating function $\mathcal{C}_0(z)$ encoding the initial data has the following asymptotic behavior in the $z \uparrow 1$ limit:

$$\mathcal{C}_0(z) = 1 + C\Gamma(1-\alpha)\,(1-z)^{\alpha-1} + \ldots \tag{4.23}$$

To establish this we first note that $\mathcal{C}_0(z=1) = N(0) = 1$. Next we differentiate the generating function $\mathcal{C}_0(z)$, use (4.22), and take the $z \uparrow 1$ limit to yield

$$\frac{d\mathcal{C}_0}{dz} = \sum_{k\geq 1} kc_k(0)z^{k-1} \simeq C\int_0^\infty dk\,k^{1-\alpha}\,e^{k\ln z} \simeq C\int_0^\infty dk\,k^{1-\alpha}\,e^{-k(1-z)} = C\Gamma(2-\alpha)(1-z)^{\alpha-2} \tag{4.24}$$

Integrating this we indeed obtain (4.23). Using (4.23), taking the limits $t\to\infty$ and $z\uparrow 1$, and keeping only leading terms we transform (4.20) into

$$\mathcal{C}(z,t) = \frac{t^{-1}}{1+Dt(1-z)^{\alpha-1}}\,, \qquad D = -C\Gamma(1-\alpha)$$

This expression is compatible with the scaling form

$$c_k \simeq t^{-1}(Dt)^{-1/(\alpha-1)}\,f_\alpha(w)\,, \quad w = k/(Dt)^{1/(\alpha-1)} \tag{4.25}$$

of the mass distribution. Indeed, a calculation analogous to that of (4.24) gives

$$\mathcal{C}(z,t) = \sum_{k\geq 1} c_k(t) z^k \simeq t^{-1} \int_0^\infty dw\, f_\alpha(w)\, e^{-ws} = \frac{t^{-1}}{1+s^{\alpha-1}}$$

where $s = (Dt)^{1/(\alpha-1)}(1-z)$. Therefore the scaling function $f_\alpha(w)$ is the inverse Laplace transform of $(1+s^{\alpha-1})^{-1}$. The small and large s asymptotics of the Laplace transform imply corresponding large and small w behaviors of the scaling function

$$f_\alpha(w) \simeq \begin{cases} \frac{1}{\Gamma(\alpha-1)}\, w^{-(2-\alpha)} & w \to 0 \\ -\frac{1}{\Gamma(1-\alpha)}\, w^{-\alpha} & w \to \infty \end{cases} \tag{4.26}$$

For instance, let $\alpha = 3/2$. In this situation, the scaling function has an explicit expression

$$f_{3/2}(w) = \frac{1}{\sqrt{\pi w}} - e^w \mathrm{erfc}\left(\sqrt{w}\right)$$

in terms of the error function.

In the marginal case $\alpha = 2$, the analysis is essentially the same, e.g. (4.23) becomes

$$\mathcal{C}_0(z) = 1 + C\,(1-z)\ln(1-z) + \ldots$$

The scaling solution is a pure exponential

$$c_k(t) = t^{-2}(C\ln t)^{-1}\, \exp\left[-\frac{k}{Ct\ln t}\right]$$

and only the appearance of logarithms in the scaling variable and in the prefactor shows distinguishes from the standard scaling behavior valid for all $\alpha > 2$.

Overall, for $\alpha \geq 2$ the scaling function is universal and pure exponential, while for each $1 < \alpha < 2$ there is a specific scaling function $f_\alpha(w)$ that is the inverse Laplace transform of $(1+s^{\alpha-1})^{-1}$. These scaling functions are still universal in that they are independent on details of the initial data (small mass behavior, the amplitude C, *etc.*), the only relevant feature is the magnitude of the decay exponent α.

Product Kernel, $K_{ij} = ij$

When the rate of aggregation is a sufficiently increasing function of masses of the reacting clusters, gelation can occur. This is the phenomenon in which a finite fraction of the total mass of the system condenses into an infinite-mass cluster in a finite time — such as the setting of Jello. The product kernel represents and exactly soluble example of this spectacular feature. Beyond the *gelation time*, the system divides into two phases: the *gel*, or the infinite cluster, and the remaining *sol* of finite clusters whose total mass decreases with time. The product kernel naturally arises for monomers that consist of f-functional reactive endgroups (Fig. 4.3). When two monomers merge, the resulting dimer has $2f-2$ reactive endgroups, a trimer has $3f-4$ endgroups, and a k-mer has $kf - 2(k-1) = (f-2)k + 2$ endgroups. If all endgroups are equally reactive, the reaction rate between two clusters equals the product of the number of endgroups. Thus

$$K_{ij} = [(f-2)i+2][(f-2)j+2] = (f-2)^2 ij + 2(f-2)(i+j) + 4. \tag{4.27}$$

The case $f = 2$ corresponds to linear polymers, for which K_{ij} is constant, while the product kernel arises for $f \to \infty$. For finite $f > 2$, the kernel is a linear combination of the constant, product, and sum kernels.

Let us focus on the product kernel, $K_{ij} = ij$, for which the master equations are

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k} ij\; c_i c_j - k c_k \sum_i i\, c_i = \frac{1}{2}\sum_{i+j=k} ij\; c_i c_j - k c_k. \tag{4.28}$$

There is already a subtlety in merely writing the master equations. In the loss term, a k-mer disappears if it reacts with a cluster of *any* size, include an infinite-mass gel, if it exists. Thus the sum in the loss term, $\sum k c_k$, includes all finite clusters *and* the gel, so that $\sum k c_k = 1$. However, when we compute the rate equation for the first moment by summing the master equations, the resulting sum, $\sum k\dot{c}_k$, is not necessarily conserved because the term associated with the infinite-mass cluster is not included in this sum. We will return to this point shortly.

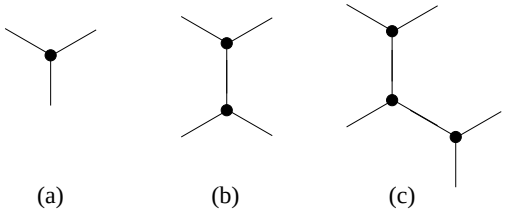


Figure 4.3: Small k-mers of 3-functional units. (a) Monomer. (b) Dimer. (c) Trimer.

Moments

If we were unaware of the existence of a singularity, the quickest way to detect that something odd may occur is from the behavior of the moments. Summing equations (4.28) we get

$$\frac{dN}{dt} = \frac{1}{2}\sum_{i,j} ic_i\, jc_j - \sum_k kc_k = \frac{1}{2} - 1 = -\frac{1}{2} \tag{4.29}$$

The solution $N(t) = 1 - \frac{t}{2}$ vanishes at $t = 2$ and becomes negative when $t > 2$. This pathology is the sign that a gel appears at some t_g (that is less than 2), after which $\sum kc_k = 1$ is no longer valid. Thus equation (4.29) must be modified when $t > t_g$; we will see that the right modification ensures that the cluster density remains positive.

The above argument predict only the upper bound $t_g < 2$, but the behavior of higher moments suggests that the gelation time $t_g = 1$. Consider the second moment M_2, which evolves as

$$\frac{dM_2}{dt} = \sum_k k^2\, \dot{c}_k = \frac{1}{2}\sum_i\sum_j \left[(i+j)^2(ic_i)(jc_j) - k^3 c_k\right] \tag{4.30}$$

$$= \frac{1}{2}\sum_i\sum_j \left[(i^3c_i)(jc_j) + (ic_i)(j^3c_j) + 2(i^2c_i)(j^2c_j) - k^2c_k\right] \tag{4.31}$$

$$= M_2^2. \tag{4.32}$$

Solving this equation subject to $M_2(0) = 1$ we obtain $M_2(t) = (1-t)^{-1}$. (For a general initial condition, the solution of (4.30) is singular when $t = 1/M_2(0)$.) The singularity is the sign of gelation, and suggests that gelation occurs at $t_g = 1$. However, is it possible that the third moment diverges earlier, so that gelation must have occurred earlier? The answer is *no*. Indeed, writing the rate equation for the third moment:

$$\frac{dM_3}{dt} = \frac{1}{2}\sum_{i,j}(i+j)^3 ic_i\, jc_j - \sum_k k^4 c_k\,, = 3M_3M_2 \tag{4.33}$$

and solving subject to $M_3(0) = 1$ we obtain $M_3(t) = (1-t)^{-3}$. Similarly,

$$\frac{dM_4}{dt} = \frac{1}{2}\sum_{i,j}(i+j)^4 ic_i\, jc_j - \sum_k k^5 c_k = 4M_4M_2 + 3M_3^2\,, \tag{4.34}$$

whose solution is $M_4(t) = (1+2t)(1-t)^{-5}$. Using induction one may verify that all moments diverge at $t_g = 1$. However, the moment method does not allow one to probe the moments (and the mass distribution) beyond the gel point. For this more complete analysis we need the generating function technique.

Generating function approach

To solve Eqs. (4.28), it is convenient to use the *exponential* generating function $\mathcal{E}(y,t) \equiv \sum_k k\, c_k(t)\, e^{yk}$. This generating function encodes the sequence kc_k instead of the sequence c_k and makes the ensuing analysis

slightly simpler. To determine the governing equation for $\mathcal{E}$ we multiply the master equation for each $\dot{c}_k$ by $k\,e^{yk}$ and sum over k to obtain

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial t} &= \frac{1}{2}\sum_{i,j}(i+j)ij\,c_i\,c_j\,e^{yk} - \sum_k k^2 c_k\,e^{yk} \\ &= \frac{1}{2}\sum_i i^2 c_i\,e^{yi}\sum_j jc_j\,e^{yj} + \frac{1}{2}\sum_i c_i\,e^{yi}\sum_j j^2 c_j\,e^{yj} - \sum_k k^2 c_k\,e^{yk} \\ &= (\mathcal{E}-1)\frac{\partial \mathcal{E}}{\partial y}. \end{aligned} \tag{4.35}$$

This is the Burgers equation — the simplest non-linear hyperbolic equation. The salient feature of the Burgers equation is that it describes the development of shock waves. The appearance of a gel in product-kernel aggregation is closely related to this appearance of a shock wave.

Equations such as (4.35) can be transformed into a linear equation by the hodograph transformation[2] that interchanges the role of the dependent and independent variables. We first write partial derivatives in terms of the Jacobian:

$$\frac{\partial \mathcal{E}}{\partial t} = \frac{\partial(\mathcal{E},y)}{\partial(t,y)}.$$

Then Eq. (4.35) for the generating function can be re-written as:

$$\frac{\partial \mathcal{E}}{\partial t} = \frac{\partial(\mathcal{E},y)}{\partial(t,y)} = (\mathcal{E}-1)\frac{\partial \mathcal{E}}{\partial y} = (\mathcal{E}-1)\frac{\partial(\mathcal{E},t)}{\partial(y,t)}.$$

Now we cancel out the common factor in the denominator to obtain the implicit, but linear equation for the generating function

$$\frac{\partial(\mathcal{E},y)}{\partial(\mathcal{E},t)} = \frac{\partial y}{\partial t}\Big|_{\mathcal{E}} = 1-\mathcal{E}. \tag{4.36}$$

The solution is simply $y = (1-\mathcal{E})t + f(\mathcal{E})$, where $f(\mathcal{E})$ is determined from the initial condition. For the monomer-only initial condition, the initial generating function is $\mathcal{E}(t=0) = \sum k c_k\,e^{yk}|_{t=0} = e^y$, or $y(t=0) = f(\mathcal{E}) = \ln \mathcal{E}$. Hence we arrive at the implicit solution

$$\mathcal{E}\,e^{-\mathcal{E}t} = e^{y-t}. \tag{4.37}$$

The generating function itself is obtained by the Lagrange inversion formula (see highlight below). Identifying $y = \mathcal{E}t$ and $x = te^{y-t}$ in Eq. (4.40) immediately gives

$$\mathcal{E}t = \sum_{k\geq 1}\frac{k^{k-1}}{k!}\,t^k\,e^{-kt}\,e^{yk}\,.$$

Since the density c_k equals the k^{th} term in the series expansion of $\mathcal{E}$ divided by k, we obtain the remarkably simple result

$$c_k(t) = \frac{k^{k-2}}{k!}\,t^{k-1}\,e^{-kt}\,. \tag{4.38}$$

[2] An alternative is to write $y = y(h,t)$, compute $dy = y_h dh + y_t dt$ and then relate the derivatives when $dy = 0$.

Lagrange inversion

Given a function $x = f(y)$, with $x \sim y$ for small y, what is the power-series representation of the inverse function $y(x) = \sum_{n\geq 1} A_n x^n$? The coefficients A_n are given by the Lagrange inversion formula. Formally, the coefficients A_n may be obtained by a contour integration around a small circle centered at the origin:

$$A_n = \frac{1}{2\pi i}\oint \frac{y}{x^{n+1}}\,dx = \frac{1}{2\pi i}\oint \frac{y}{x^{n+1}}\,\frac{dx}{dy}\,dy = \frac{1}{2\pi i}\oint \frac{y}{f(y)^{n+1}}\,f'(y)\,dy. \tag{4.39}$$

The crucial step is to transform from integrating over x to integrating over y. The transformed contour is also a small circle about the origin since y and x are proportional to each other near the origin.

Let's apply this inversion formula to $f(y) = y\,e^{-y} = x$. From Eq. (4.39) and using $\frac{dx}{dy} = (1-y)\,e^{-y}$, we have

$$A_n = \frac{1}{2\pi i}\oint \frac{y}{(y\,e^{-y})^{n+1}}\,(1-y)\,e^{-y}\,dy = \frac{1}{2\pi i}\oint \frac{1-y}{y^n}\,e^{ny}\,dy$$

To find the residue we simply expand the exponential in a power series and then read off the coefficient of $\frac{1}{y}$ in the integral. Thus

$$A_n = \frac{1}{2\pi i}\oint \sum_{k=0}^{\infty}\frac{n^k}{k!}\left(y^{k-n} - y^{k+1-n}\right)dy = \frac{n^{n-1}}{(n-1)!} - \frac{n^{n-2}}{(n-2)!} = \frac{n^{n-1}}{n!},$$

so that the series representation of the inverse function $y(x)$ is

$$y = \sum_{n\geq 1}\frac{n^{n-1}}{n!}\,x^n. \tag{4.40}$$

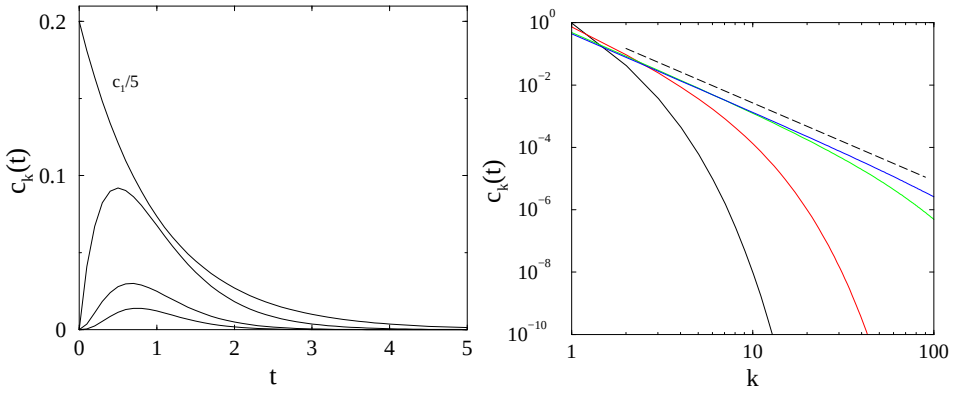


Figure 4.4: Left: Cluster concentrations $c_k(t)$ versus time for the product kernel for $k = 1, 2, 3, 4$ (top to bottom, with c_1 divided by 5). Right: $c_k(t)$ versus k for $t = 0.1, 0.4, 0.8$, and 0.9 on a double logarithmic scale (upper left to lower right). The dashed line has slope $-5/2$.

For the asymptotic behavior of this distribution, Stirling's approximation gives

$$c_k(t) \simeq \frac{k^{k-2}}{\sqrt{2\pi k}}\left(\frac{e}{k}\right)^k t^{k-1}\,e^{-kt} \underset{t=1-\epsilon}{\longrightarrow} \frac{e^{-k\epsilon^2/2}}{\sqrt{2\pi}\,k^{5/2}}, \tag{4.41}$$

where we have approximated $e^{-k(t-\ln t-1)}$ by $e^{-k(1-t)^2/2}$ for $t \to 1$. For $t \neq 1$, the mass distribution c_k decreases exponentially with k. At the gelation time $t = t_g = 1$, however, the mass distribution has a

power-law tail that is a precursor of the singularity where an infinite-mass cluster first appears. Beyond t_g, the cluster population naturally divides into the sol and the gel phases. Near the gelation time, (4.41) gives the scaling form for the mass distribution

$$c_k(t) \simeq s^{-5/2}\Phi(k/s) \qquad \text{with} \quad \Phi(z) = \frac{1}{\sqrt{2\pi}}\,\frac{e^{-z/2}}{z^{5/2}}, \tag{4.42}$$

in which the characteristic mass is $s = (1-t)^{-2}$.

The behavior of the moments $M_n = \sum_{k\geq 1} k^n c_k$ of the mass distribution cleanly illustrates what is happening near the gelation transition. The most dramatic behavior occurs for the first moment $M_1 = \sum k c_k$ — ostensibly the total mass — which is conserved only for $t \leq t_g$. Beyond t_g, the sum in M_1 accounts for the mass of *finite* clusters only, while the contribution of an infinite-mass cluster is excluded. Thus $g \equiv 1 - M_1$ gives the fraction of the total mass that belongs to the infinite cluster or the gel. To find g, we substitute $y = 0$ in the implicit equation (4.37) for the generating function and then use $g \equiv 1 - M_1$ to give

$$g = 1 - e^{-gt}\,. \tag{4.43}$$

This equation always admits a trivial solution $g = 0$. For $t > 1$, however, there is an additional non-trivial solution in which the gel has a non-zero mass. While Eq. (4.43) is not analytically soluble, the limiting behaviors of Eq. (4.43) can be obtained perturbatively. Just past the gelation time, we write $t = 1 + \delta$ and expand (4.43) for small δ, while for $t \to \infty$, we write $g = 1 - \epsilon$ and expand for small ϵ. These give

$$g = \begin{cases} 0 & \text{for} \quad t < 1 \\ 2(t-1) - 8(t-1)^2/3 + \ldots & \text{for} \quad t \downarrow 1 \\ 1 - e^{-t} - te^{-2t} + \ldots & \text{for} \quad t \to \infty. \end{cases} \tag{4.44}$$

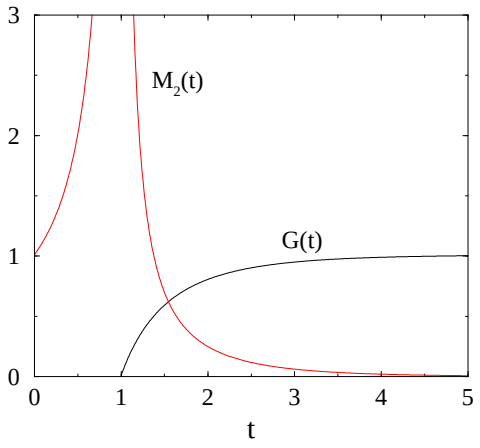


Figure 4.5: Time dependence of the mass of gel phase and the mean mass of the finite clusters.

Similarly we may obtain rate equations for all moments that are valid in the post-gel regime. For example, the zeroth moment, or the density of finite clusters, [3] $M_0 \equiv N = \sum c_k$, obeys

$$\begin{aligned} \frac{dM_0}{dt} &= \frac{1}{2}\sum_{i,j} ic_i\, jc_j - \sum_k kc_k \\ &= \frac{1}{2}\,(1-g)^2 - (1-g) = \frac{g^2-1}{2}. \end{aligned} \tag{4.45}$$

[3]For the zeroth moment, the distinction between the density of all clusters and finite clusters is immaterial since there is only a single infinite cluster.

This equation confirms our expectation that (4.29) is correct only in the pre-gel regime. Using the results of (4.44) for g, we obtain

$$M_0 = \begin{cases} 1 - t/2 & \text{for} \quad t \leq 1; \\ 1 - t/2 + 2(t-1)^3/3 + \ldots & \text{for} \quad t \downarrow 1; \\ e^{-t} + (t/2)\, e^{-2t} + \ldots & \text{for} \quad t \to \infty. \end{cases} \tag{4.46}$$

Similarly, the rate equation for M_2 is

$$\begin{aligned} \dot{M}_2 &= \frac{1}{2}\sum_{i,j}(i+j)^2\, i\, j\, c_i\, c_j - \sum_{i,k} k^3\, c_k = \sum_{i,j}(i^3\, c_i\, j c_j + i^2\, c_i\, j^2\, c_j) - \sum_{i,k} k^3\, c_k \\ &= M_2^2 - M_3 g \end{aligned}$$

Before the gel point we recover the already known solution $M_2(t) = (1-t)^{-1}$. For $t > t_g$ the equation for M_2 involves g and M_3 which are not known explicitly. Therefore there is no explicit expression for M_2 and indeed for higher moments, in the post-gel regime.

Interestingly, the higher moments can be expressed in terms of g. First, we note that the moments for the population of finite clusters are just the derivatives of the generating function $\mathcal{E}$:

$$M_n = \left.\frac{\partial^{n-1}\mathcal{E}}{\partial y^{n-1}}\right|_{y=0}.$$

Let us consider the second moment. We take the logarithm of Eq. (4.37), differentiate with respect to y, and set $y = 0$ to give

$$M_2(t) = \left[\frac{1}{\mathcal{E}(y=0,t)} - t\right]^{-1}, \tag{4.47}$$

with $\mathcal{E}(y=0,t) = 1$ in the sol phase and $\mathcal{E}(y=0,t) = 1 - g = e^{-gt}$ [see Eq. (4.43)] in the gel phase. Therefore,

$$M_2(t) = \begin{cases} (1-t)^{-1} & \text{for} \quad t < 1; \\ (e^{gt} - t)^{-1} & \text{for} \quad t > 1. \end{cases} \tag{4.48}$$

For $t \to t_g$ from below, the second moment grows rapidly with time, while for large t, $M_2 \to 0$ as finite clusters are progressively engulfed by the gel (Fig. 4.5).

Sum Kernel, $K_{ij} = i + j$

For the sum kernel, the master equations now are

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k}(i+j)\, c_i c_j - c_k \sum_i (i+k)\, c_i\,. \tag{4.49}$$

Let's start by studying the moments of the mass distribution. The rate equations for the first few moments are:

$$\begin{aligned} \dot{M}_0 &= -M_1 M_0 \\ \dot{M}_2 &= 2M_1 M_2 \\ \dot{M}_3 &= 3M_1 M_3 + 3M_2^2 \\ \dot{M}_4 &= 4M_1 M_4 + 10 M_2 M_3 \end{aligned} \tag{4.50}$$

and generally the n^{th} moment obeys

$$\dot{M}_n = n M_1 M_n + \frac{1}{2}\sum_{a=2}^{n-1}\binom{n+1}{a} M_a M_{n+1-a}$$

for $n \geq 2$. Solving the moment equations for the monomer-only initial condition ($M_n(t) = 1$ for all $n \geq 0$) we obtain

$$\begin{aligned} M_0(t) &= e^{-t} \\ M_2(t) &= e^{2t} \\ M_3(t) &= 3\, e^{4t} - 2\, e^{3t} \\ M_4(t) &= 15\, e^{6t} - 20\, e^{5t} + 6\, e^{4t}\,. \end{aligned} \tag{4.51}$$

Generally for $n \geq 2$ the moment M_n involves the combination of exponentials $e^{(2n-2)t}, e^{(2n-3)t}, \ldots, e^{nt}$. All moments remain finite at all times implying that there is no gelation.

One subtle feature of sum kernel aggregation is that the definition of the typical aggregate mass $s(t)$ has an apparent ambiguity. Two natural definitions for the typical mass are:

$$s(t) = \frac{\sum_{k\geq 1} k c_k(t)}{\sum_{k\geq 1} c_k(t)} = \frac{M_1(t)}{M_0(t)}\,, \quad \text{and} \quad s(t) = \frac{\sum_{k\geq 1} k^2 c_k(t)}{\sum_{k\geq 1} k c_k(t)} = \frac{M_2(t)}{M_1(t)}\,.$$

However, if scaling holds, the ratio M_{n+1}/M_n, or even a more complicated expression $(M_{n+m}/M_n)^{1/m}$, should be equally acceptable definitions of the typical mass; for example, for constant-kernel aggregation any of the above definitions for the typical mass grow linearly with time. However, for sum-kernel aggregation we have $M_1/M_0 = e^t, (M_2/M_0)^{1/2} = e^{3t/2}, (M_3/M_0)^{1/3} \sim e^{5t/3}$, *etc.* This unconventional behavior is an outcome of a small-mass singularity in the mass distribution. However, for a scaling combination of moments that does not involve M_0, we would find a typical mass that grows as $s \sim e^{2t}$.

The master equations for the sum kernel can be solved in a number of ways. One approach is to change variables to eliminate the loss terms. This transformation is easier to see if instead of the densities c_k we consider the ratios $\psi_k \equiv c_k/N$. These ratios satisfy

$$\psi_k' = \sum_{i+j=k} i\, \psi_i \psi_j - k\, \psi_k, \tag{4.52}$$

where the prime denotes differentiation with respect to the rescaled time $\tau = \int_0^t dt'\, N(t') = 1 - e^{-t}$. Now we use $\Psi_k \equiv I\, \psi_k$ instead of ψ_k, with I being the integrating factor $I = \exp[\int_0^\tau d\tau'\, f_k(\tau')]$. Equations (4.52) now become

$$\Psi_k' = \sum_{i+j=k} i\, \Psi_i\, \Psi_j\,. \tag{4.53}$$

These equations are recurrent and can be solved one by one. For the monomer-only initial condition, $\Psi_k(0) = \delta_{k,1}$, we find $\Psi_1 = 1, \Psi_2 = \tau, \Psi_3 = \frac{3}{2}\tau^2$, *etc.* These results suggest the exponential ansatz $\Psi_k = a_k\, \tau^{k-1}$. Substituting this ansatz into (4.53), the amplitudes satisfy an algebraic recursion formula

$$(k-1)a_k = \sum_{i+j=k} i\, a_i\, a_j\,, \qquad a_1 = 1. \tag{4.54}$$

To solve this recursion, we introduce an exponential generating function $\mathcal{A}(z) = \sum_k a_k e^{kz}$ that recasts (4.54) into $d\mathcal{A}/dz = \mathcal{A}/(1-\mathcal{A})$, with solution $\mathcal{A}\, e^{-\mathcal{A}} = e^z$. This is almost the same form as the implicit solution for the generating function in the product kernel, Eq. (4.37), and it follows, after applying the Lagrange inversion formula (4.40), that the solution is $a_k = k^{k-1}/k!$. Finally, we unfold the above transformations to obtain $c_k(t)$:

$$c_k(t) = \frac{k^{k-1}}{k!}\,(1 - e^{-t})^{k-1}\, e^{-t}\, e^{-k(1-e^{-t})}\,. \tag{4.55}$$

When both k and t are large, we use Stirling's approximation and expand the logarithm to simplify the above formula to

$$c_k(t) \simeq \frac{e^{-t}}{\sqrt{2\pi k^3}}\, e^{k[\ln(1-e^{-t}) + e^{-t}]} \to \frac{e^{-t}}{\sqrt{2\pi k^3}}\, e^{-ke^{-2t}/2}\,. \tag{4.56}$$

This expression shows that the typical mass $s \sim e^{2t}$. The anomalous behavior of the zeroth moment is a consequence of the algebraic $k^{-3/2}$ prefactor, which causes a divergence in $\sum_k c_k$ at small masses — hence

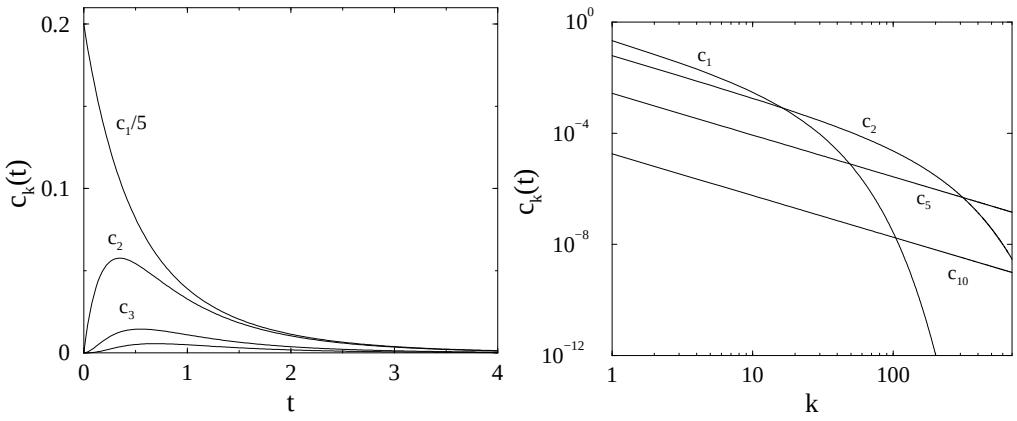


Figure 4.6: Left: Cluster concentrations $c_k(t)$ versus time for the sum kernel for $k = 1, 2, 3, 4$ (with c_1 divided by 5). Right: $c_k(t)$ versus k for $t = 1, 2, 5$, and 10 on a double logarithmic scale. For small k The straight lines have slope $-3/2$.

the zeroth moment does not probe the typical mass. For the higher moments we can use (4.56) and replace the sum by an integral to give the asymptotic behavior

$$M_n(t) \simeq \int_0^\infty dk\, k^n\, \frac{e^{-t}}{\sqrt{2\pi k^3}}\, e^{-ke^{-2t}/2} = \frac{2^{n-1}\Gamma\left(n-\frac{1}{2}\right)}{\sqrt{\pi}}\, e^{2(n-1)t} \qquad n > 0\,, \tag{4.57}$$

from which the typical size measures $(M_{n+m}/M_n)^{1/m}$ all scale as e^{2t} for any $m, n > 0$.

A quicker route to the solution for the sum kernel relies on an unexpected connection with the product kernel system. We start by writing the master equation (4.49) as

$$\dot{c}_k + c_k + k\, c_k\, N = \sum_{i+j=k} i\, c_i\, c_j\,, \tag{4.58}$$

where we use the fact that $\sum i\, c_i = 1$. Introducing the integrating factor $I = \exp[\int_0^t dt'\,(1 + k\, N(t'))]$, with $N = e^{-t}$, the quantity $\psi_k = I\, c_k$ obeys

$$\begin{aligned}
\dot{\psi}_k &= \sum_{i+j=k} i\, c_i\, c_j\; e^{[t+k(1-e^{-t})]} \\
&= e^{-t} \sum_{i+j=k} i\; c_i\, e^{[t+i(1-e^{-t})]}\, c_j\; e^{[t+j(1-e^{-t})]} \\
&= e^{-t} \sum_{i+j=k} i\, \psi_i\, \psi_j.
\end{aligned} \tag{4.59}$$

Next we define again $d\tau = e^{-t}\, dt$ to obtain $\psi' = \sum_{i+j=k} i\, \psi_i \psi_j$. Finally, by introducing the generating function $\mathcal{A}(z) = \sum \phi_k\, e^{zk}$, we recast the recursion formula for ϕ_k into $\frac{\partial \mathcal{A}}{\partial \tau} = \mathcal{A}\frac{\partial \mathcal{A}}{\partial z}$. This is the same equation of motion (4.35) for the generating function $\mathcal{E} - 1$ in the product kernel, except with the time-like variable τ instead of t. We therefore obtain a similar solution for the cluster concentrations as in the product kernel, but as a function of τ rather than t.

Other Simple Kernels

The general polymerization kernel written in Eq. (4.27)

$$K_{ij} = A + B(i + j) + Cij\,, \tag{4.60}$$

which is a linear combination of the constant, sum, and product kernels, is also soluble. Its solvability cannot be deduced by 'superposition' because the master equations are nonlinear. However, because the constant, sum, and product kernels are separately solvable by the generating function technique, and also accounting for the fact that the sum and product kernels are intimately related, it is not too surprising that the general polymerization kernel is tractable. For $C > 0$, the product term dominates and the system undergoes gelation; if $C = 0$, the sum term dominates when $B > 0$, and the typical mass grows exponentially with time. The model (4.60) is valuable because it represents most of exactly soluble aggregation models.[4]

A more challenging problem is to study kernels with more rapid aggregation, such as

$$K_{ij} = (i+j)^2 \quad \text{or} \quad K_{ij} = (ij)^2 \ . \tag{4.61}$$

Formally, these kernels are amenable to generating function techniques, but the resulting partial differential equations for the generating function are intractable because they are non-linear and higher than first order [compare with (4.36) for the product kernel]. Physically, we anticipate that the model $K_{ij} = (i+j)^2$ undergoes gelation, since it is at least as 'reactive' than the product kernel. However, even the simple moment method does help in extracting basic information. To see the cause of the trouble, consider explicit rate equations for the first few moments:

$$\begin{aligned} \dot{M}_2 &= 2M_1M_3 + 2M_2^2 \\ \dot{M}_3 &= 3M_1M_4 + 9M_2M_3 \\ \dot{M}_4 &= 4M_1M_5 + 14M_2M_4 + 10M_3^2 \\ \dot{M}_5 &= 5M_1M_6 + 20M_2M_5 + 35M_3M_4 \end{aligned} \tag{4.62}$$

Equations (4.62) are *hierarchical* — the equation for each $\dot{M}_n$ contains a *higher* moment M_{n+1}. In our previous examples, the moment equations were recurrent, and hence solvable one by one.

Hierarchical equations are generally unsolvable, and all attempts to find exact or asymptotic solutions of equations (4.62) have failed. The reason for this failure is not mathematical, but conceptual for this specific case of $K_{ij} = (i+j)^2$, because a gel appears *instantaneously* — $t_g = 0^+$! This spectacular behavior occurs not only for the kernel $K_{ij} = (i+j)^2$, but for a wide class of models with asymptotically homogeneous kernels of the form

$$K_{1,j} = K_{j,1} \sim j^\nu \quad \text{when} \quad j \to \infty \tag{4.63}$$

with ν strictly greater than 1. The applicability of such models to real aggregation processes is questionable[5] because the number of active sites on a cluster should not increase faster than its size. This implies that the ratio $K(1,j)/j$ is bounded as $j \to \infty$, leading to the upper bound $\nu \leq 1$.

4.2 Scaling

Scaling exploits the observation that the typical cluster mass changes systematically with time so that a change in time scale corresponds to a change in mass scale. This equivalence is embodied by the *scaling ansatz*, which may be written as

$$c(x,t) = \frac{1}{s^2}\, f\left(\frac{x}{s}\right).$$

Here $s = s(t)$ is the typical cluster mass, x/s is the scaled mass, and $f(x/s)$ is the scaling function. Thus the fundamental system variables are not the mass and time, but rather, the scaled mass x/s and the time. The prefactor s^{-2} in front of the scaling function enforces mass conservation: $\int x\, c(x,t)\, dx = 1$ reduces to the manifestly time-independent relation $\int u\, f(u)\, du = 1$.

[4] The remaining exactly solved models include the exponential kernel $K_{ij} = 2 - q^i - q^j$ with $0 < q < 1$, that interpolates between the constant kernel, $K_{ij} = 2$ when $q = 0$, and the sum kernel, $K_{ij} \simeq (1-q)(i+j)$ when $q \to 1-0$. The other exactly solved model is the parity-dependent kernel where K_{ij} takes 3 distinct values depending on whether i, j are both even, both odd, or one index is even and the other is odd.

[5] Nevertheless, kernels with $\nu > 1$ have been proposed in various contexts, *e.g.*, for gravitationally attracting particles.

There are several reasons why scaling plays a central role in numerous non-equilibrium phenomena. The chief reason, of course, is that it tremendously simplifies and condenses the description — a function of one variable is infinitely simpler than a function of two variables. On a more technical level, scaling provides the simplest route to the asymptotic solution of the master equations, especially for problems where exact solutions are difficult or impossible to obtain. This simplification arises because the scaling ansatz separates a two-variable master equation into two simpler single-variable systems that can be analyzed individually. Further, a scaling solution is independent of the initial conditions[6] and thus is automatically constructed to focus on the interesting asymptotic behavior. Finally, scaling gives universal information aspects of the asymptotic mass distribution in terms of generic features of the reaction kernel and it provides a robust classification of the solutions to the master equations for many non-equilibrium processes; we will see this approach in action in many of the later chapters.

Before we can apply scaling, we need to settle on the "right" definition for the typical mass. From the scaling ansatz, the n^{th} moment of the mass distribution is

$$M_n = \int x^n \, \frac{1}{s^2} \, f(x/s)\, dx \sim s^{n-1}.$$

Hence for any value of n, the ratio M_{n+1}/M_n is proportional to s. Consequently, either $1/M_0$ or M_2 (where we set $M_1 = 1$) are good measures of the typical mass, as long as the cluster mass distribution itself is not too singular. When scaling holds, we can define the typical mass to best suit the situation.

We also need basic information about the matrix of reaction rates K_{ij} to determine the consequences of scaling. It turns out that only two features of this matrix determine the asymptotic properties of the mass distribution. The first is the homogeneity index λ, defined by

$$K_{ai,aj} \sim a^\lambda K_{ij},$$

that gives the overall mass dependence of the reaction rate. The second is the index ν, defined by

$$K_{1,j} = K_{j,1} \sim j^\nu,$$

that characterizes the relative importance of reactions between clusters of similar masses and disparate masses. For example, the constant kernel is characterized by $(\lambda, \nu) = (0, 0)$ and the product kernel by $(\lambda, \nu) = (2, 1)$. An important example is the "Brownian" kernel (4.4), the reaction rate for spherical aggregates that undergo Brownian motion. For this kernel, $(\lambda, \nu) = (0, 1/3)$.

The role of the indices λ and ν may be best appreciated by considering the following pictorial representation of the reaction matrix

$$K_{ij} = \begin{pmatrix} SS & \cdots & SL & \cdots \\ \vdots & \ddots & \cdots & \cdots \\ LS & \vdots & LL & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

The meta-entries SS, $SL(= LS)$, and LL denote the reaction rates of small clusters with other small clusters, large-small interactions, and large-large interactions, respectively. The exactly-soluble examples discussed above are archetypes of three distinct universality classes with the following general behavior:

- **Type I:** $LL \gg LS, SS$, corresponding to $\lambda > \nu$. Because of the high reactivity of large clusters they quickly disappear, while small clusters tend to persist. The result is a cluster mass distribution that decays monotonically with mass. The product kernel typifies this type of system.

- **Type II:** all three reactions are of the same order. This marginal class contains the simplest soluble case of $K_{ij} = 1$. However the asymptotic behavior of this class is sensitive to details of the reaction rates.

[6] Apart from a trivial dependence on a few basic features, like the total mass; see *e.g.*, Eq. (4.21). Exceptions to the rule that initial conditions do not affect the scaling function arise if the initial data have slowly decaying tails; see example 3 in Sec. 4.1. Such initial data are considered pathological.

- **Type III:** $LS \gg LL, SS$, or $\lambda < \nu$. As the reaction develops, small clusters are quickly removed from the system because of the dominance of large-small interactions. Thus the system has a dearth of small clusters, leading to a peaked mass distribution.

Let's now apply scaling to determine basic features of the cluster mass distribution. In the continuum limit, the master equations for aggregation are

$$\dot{c}(x,t) = \frac{1}{2}\int_0^x dy\, K(y, x-y)\, c(y,t)\, c(x-y,t) - \int_0^\infty dy\, K(x,y)\, c(x,t)\, c(y,t) \tag{4.64}$$

Using homogeneity, $K(ax, ay) = a^\lambda K(x,y)$, and substituting the scaling form $c(x,t) = s^{-2}\, f(x/s)$ into the master equation (4.64), the left-hand side becomes

$$\dot{c}(x,t) = -\frac{\dot{s}}{s^3}\left[2f(u) + uf'(u)\right],$$

where $u = x/s$, while the right hand side is $s^{\lambda-3}\mathcal{K}(u)$, where

$$\mathcal{K}(u) = \frac{1}{2}\int_0^u dv\, K(v, u-v)\, f(v)\, f(u-v) - \int_0^\infty dv\, K(u,v)\, f(u)\, f(v)\,, \tag{4.65}$$

with $v = y/s$. Equating and re-arranging, the dependences on time and on the scaled mass u separate as

$$\frac{\dot{s}(t)}{s(t)^\lambda} = -\frac{\mathcal{K}(u)}{2f(u) + uf'(u)} \equiv \Lambda. \tag{4.66}$$

The left-hand side is a function of time only while the right-hand side is a function of u only, so that they are both separately equal to a constant — the separation constant Λ. (Actually, there is a time dependence hidden in ϵ on the right-hand side that disappears as long as the integrals converge at their lower limits.) This variable separation is a primary simplifying feature of the scaling ansatz.

The time dependence of the typical mass is determined from $\dot{s} = \Lambda s^\lambda$ and gives three different behaviors:

$$s(t) \sim \begin{cases} t^{1/(1-\lambda)} \equiv t^z & \lambda < 1; \\ e^{\Lambda t} & \lambda = 1; \\ (t_g - t)^{-1} & 1 < \lambda \leq 2. \end{cases} \tag{4.67}$$

For non-gelling systems the time dependence of the typical mass is primarily determined by the homogeneity index λ; other features of the reaction rate such as the second homogeneity index ν affect only details. For instance, in the growth law $s(t) \simeq At^{1/(1-\lambda)}$, the amplitude A depends on details of the reaction kernel while the growth exponent $1/(1-\lambda)$ depends only on λ.

The time dependence (4.67) can also be obtained from the following heuristic argument. Assuming scaling with a typical cluster mass s at time t, the corresponding cluster density is of the order of $1/s$. Consider a time increment Δt during which all clusters react, so that the typical mass increases by $\Delta s \approx s$. This time increment is the inverse of an overall reaction rate. In turn, this rate is proportional to the reaction kernel $K(s,s) \sim s^\lambda$ and the concentration $1/s$. Hence $\frac{\Delta s}{\Delta t} \sim s \times (s^\lambda/s)$ which then reproduces (4.67).

We see that non-gelling systems correspond to $\lambda \leq 1$, while for gelling systems $1 < \lambda \leq 2$. We should keep in mind, of course, that the master equations are ill-posed if $\nu > 1$, as instantaneous gelation occurs in this case; our consideration of (homogeneous) aggregation kernels tacitly assumes that $\nu \leq 1$.

The dependence of the scaling function f on the scaled mass u is governed by the u-dependent part of (4.66),

$$2f(u) + uf'(u) + \Lambda^{-1}\mathcal{K}(u) = 0 \tag{4.68}$$

with $\mathcal{K}(u)$ given by (4.65). The non-linear integro-differential equation (4.68) is complicated, and the full understanding of the behavior of the scaling function $f(u)$ is still lacking. It is certainly impossible to solve (4.68) for an arbitrary kernel, so 'understanding' refers to qualitative features: asymptotic behaviors, justifying the classification to type I, II, and III kernels, etc. For instance, it has been shown that when the scaled mass is large, $u \gg 1$, the scaling function exhibits a fairly simple behavior, namely it is exponential

$f(u) \propto e^{-au}$; the detailed behavior of the kernel affects only the prefactors. The behavior of the scaling function when the scaled mass is small is much less robust. Many empirical results indicate that $f(u) \sim u^{-\tau}$ for $u \ll 1$. As a corollary, the time dependence of the density of small-mass clusters is given by

$$c_k \sim \frac{1}{s^2}\left(\frac{k}{s}\right)^{-\tau} \sim k^{-\tau}\, t^{-(2-\tau)z} \tag{4.69}$$

The exponent τ apparently depends on the detailed properties of the reaction kernel. A heuristic approach is to *assume* that the behavior of monomers represents the $u \to 0$ limit. The master equation for the monomer density is

$$\dot{c}_1 = -c_1 \sum_{j\geq 1} K_{1j}\, c_j \tag{4.70}$$

Since $K_{1j} \sim j^{\nu}$ and $c_j \sim j^{-\tau}$, the sum on the right-hand side of (4.70) converges if $\nu - \tau + 1 < 0$, and in this case the first first term provides a good estimate for the sum. Therefore $\dot{c}_1 \approx -K_{11}\, c_1^2$, leading to $c_1 \sim 1/t$. Matching this time dependence with that given in Eq. (4.69), we deduce the exponent relation $\tau = 1 + \lambda$. This power-law tail applies for $\nu - \tau + 1 = \nu - \lambda < 0$, that is, for Type I kernels.

4.3 Aggregation with Input

Many physical realizations of aggregation do not occur in a closed system, but instead a steady input helps drive the reaction. Examples of aggregation with input are diverse, and range from chemical processing in continuously-stirred tank reactor, to the distribution of star masses in the galaxy. In all cases, the interplay between input and aggregation leads to many new phenomena. Here we consider the simplest situation of a constant input that begins at $t = 0$ and we limit ourselves to the situation of monomer input. Because the asymptotic behavior is again independent of initial conditions, we also consider only the simplest case of an initially empty system, $c_k(0) = 0$.

Constant kernel

The evolution of the mass distribution is now described by the master equation

$$\dot{c}_k = \sum_{i+j=k} c_i c_j - 2c_k N + \delta_{k,1}\,. \tag{4.71}$$

The total density satisfies $\dot{N} = -N^2 + 1$ whose solution is, for an initially empty system,

$$N(t) = \tanh t. \tag{4.72}$$

Hence the total density initially grows linearly with time but eventually saturates to 1.

The individual densities can be in principle found by solving the master equations one by one. However, again the generating function approach is a much more potent tool. We introduce the generating function $\mathcal{C}(z,t) = \sum_{k\geq 1} c_k(t) z^k$ to recast the master equations (4.71) into the differential equation [compare with Eq. (4.18)]

$$\dot{\mathcal{C}}(z,t) = \mathcal{C}(z,t)^2 - 2\mathcal{C}(z,t)N(t) + z\,. \tag{4.73}$$

As in Eq. (4.18), it is convenient to define $\mathcal{C}_- = \mathcal{C} - N$ that then satisfies the closed equation $\dot{\mathcal{C}}_- = \mathcal{C}_-^2 + (z-1)$. Solving for $\mathcal{C}_-$, we obtain

$$\mathcal{C}(z,t) = N(t) - \sqrt{1-z}\,\tanh\left(t\sqrt{1-z}\right) \;\longrightarrow 1 - \sqrt{1-z} \qquad t \to \infty. \tag{4.74}$$

The generating function at infinite time can be inverted by expanding $\sqrt{1-z}$ in a power series in z

$$\begin{aligned}\sqrt{1-z} &= 1 + \tfrac{1}{2}(-z) + \tfrac{1}{2}\left(-\tfrac{1}{2}\right)\frac{(-z)^2}{2!} + \tfrac{1}{2}\left(-\tfrac{1}{2}\right)\left(-\tfrac{3}{2}\right)\frac{(-z)^3}{3!} + \tfrac{1}{2}\left(-\tfrac{1}{2}\right)\left(-\tfrac{3}{2}\right)\left(-\tfrac{5}{2}\right)\frac{(-z)^4}{4!} + \ldots \\ &= 1 - \frac{\Gamma(k-\frac{1}{2})}{2\Gamma(\frac{1}{2})}\frac{z^k}{\Gamma(k+1)}\,,\end{aligned}$$

In deriving of this expression we use the identity

$$a(a+1)\ldots(a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)} \ ,$$

that follows from the basic gamma function identity $a\Gamma(a) = \Gamma(a+1)$. Finally, using $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, we obtain[7]

$$c_k = \frac{1}{\sqrt{4\pi}} \ \frac{\Gamma(k-\frac{1}{2})}{\Gamma(k+1)} \, . \tag{4.75}$$

For the asymptotic behavior, we use the handy asymptotic relation for $k \gg 1$,

$$\frac{\Gamma(k+a)}{\Gamma(k+b)} \simeq k^{a-b} \ ,$$

to obtain the tail of the steady-state mass distribution

$$c_k \simeq \frac{1}{\sqrt{4\pi}} \ \frac{1}{k^{3/2}} \qquad\qquad k \gg 1 \, . \tag{4.76}$$

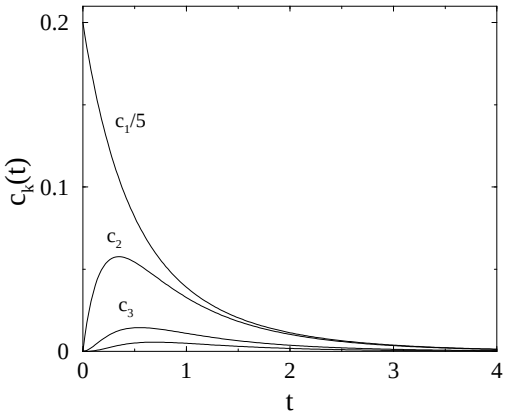


Figure 4.7: Left: Cluster mass distribution for constant-kernel aggregation with a steady monomer input. **figure not done yet.**

The steady-state mass distribution is heavy-tailed, as it must to produce a divergent mass density. At finite time, however, mass conservation requires that $\sum k c_k(t) = t$, so that the tail of the mass distribution must deviate from (4.76). However, we can give a simple qualitative argument that determines the time-dependent behavior: for sufficiently small masses $k \ll k_*$ the mass distribution $c_k(t)$ is very close to stationary form (4.76), while for $k \gg k_*$ the mass distribution is essentially zero. We determine the crossover mass k_* by requiring that the total mass in the system equals the elapsed time. Thus

$$t = \sum_{k=1}^{\infty} k c_k(t) \approx \sum_{k=1}^{k_*} k c_k \sim \sum_{k=1}^{k_*} k^{-1/2} \sim k_*^{1/2} \, , \tag{4.77}$$

leading to $k_* \sim t^2$. Thus the bulk of the population follows the steady-state power-law distribution (4.75) whose leading edge is cut off at $k_* \sim t^2$ (Fig. 4.7).

This qualitative picture can be sharpened by an exact analysis. To extract the densities from the generating function (4.74) we substitute the series representation

$$\pi \, \tanh(\pi x) = \sum_{n=-\infty}^{\infty} \frac{x}{x^2 + \left(n+\frac{1}{2}\right)^2}$$

[7] In this section we write c_k instead of $c_k(\infty)$; whenever we treat a non-steady mass distribution, we write $c_k(t)$.

into (4.74) and expand in powers of z to yield

$$c_k(t) = \frac{1}{t^3}\sum_{n=-\infty}^{\infty}\left(n+\frac{1}{2}\right)^2\pi^2\left[1+\left(n+\frac{1}{2}\right)^2\frac{\pi^2}{t^2}\right]^{-k-1} . \tag{4.78}$$

In the long-time limit, we replace the sum on the right-hand side of (4.78) by the integral over the variable $x = \left(\frac{1}{2}+n\right)\frac{\pi}{t}$. When $k \ll t^2$, we obtain

$$c_k \simeq \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x^2\,dx}{(1+x^2)^{k+1}} ,$$

and computing the integral recovers (4.75). On the other hand, when k and $t \to \infty$ such that $\kappa = k/t^2$ remains finite, (4.78) simplifies to

$$c_k(t) = \frac{1}{t^3}\sum_{n=-\infty}^{\infty}\left(n+\frac{1}{2}\right)^2\pi^2\exp\left[-\left(n+\frac{1}{2}\right)^2\pi^2\,\kappa\right] . \tag{4.79}$$

When the mass $k \gg k_*$, we may keep only the first term in (4.79) to give the leading asymptotic behavior of the mass density

$$c_k(t) = \frac{\pi^2}{4\,t^3}\;e^{-\pi^2\kappa/4} , \tag{4.80}$$

so that large-mass clusters are exceedingly rare as shown in Fig. 4.7.

Generalized sum kernel

In epitaxial surface growth, a constant flux of atoms impinges on a clean surface. In a suitable temperature range, these incident atoms adsorb irreversibly and then diffuse freely on the surface. Mobile adatoms can then merge to form dimers, trimers, *etc.*, and islands of all sizes can diffuse on the surface and continue to aggregate. The growth of islands is therefore driven by the combined effects of irreversible aggregation, island diffusion, and steady monomer input.

The mechanism underlying island diffusion is that adatoms on the edge of an island can hop to neighboring sites on the periphery (Fig. 4.8). Consider a large island of linear size R. In a time interval $\Delta t \sim R^2$, an adatom on the edge typically explores the entire island boundary,[8] and hence in a time interval Δt, each periphery adatom moves by a distance R from its original position. This diffusion of periphery adatoms leads to an effective center-of-mass displacement of the island $\delta x \sim R/R^2 \sim R^{-1}$. If each periphery adatom diffuses independently, the total center-of-mass displacement Δx will be the sum of R independent identically distributed random variables. Consequently $\Delta x \sim \sqrt{R(\delta x)^2} \sim R^{-1/2}$. Thus the effective diffusion coefficient of an island of linear dimension R scales as $D_R \sim (\Delta x)^2/\Delta t \sim R^{-3}$. Since the mass of an island of linear size R scales $k \sim R^2$, we conclude that the effective diffusion coefficient of an island of mass k scales as

$$D_k \sim k^{-3/2} . \tag{4.81}$$

The *generalized sum kernel* now arises by using the Smoluchowski formula $K_{ij} \propto (D_i + D_j)(R_i + R_j)^{d-2}$ in the diffusion-controlled limit [see Eq. (2.50)] for the aggregation rate of an i-mer and j-mer on the surface. Here R_j is the radius of a j-mer and D_j is its diffusion coefficient. For a two-dimensional surface[9], the kernel reduces to $K_{ij} \sim D_i + D_j$. For the purposes of this discussion, we write $D_j \propto j^\nu$, where ν is the mobility exponent that equals 3/2 for periphery diffusion, but could be different for other surface relaxation mechanisms.

Let's now determine the steady-state solution to the master equations for the generalized sum kernel with steady monomer input. The master equations are

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k}(i^\nu + j^\nu)\,c_i c_j - c_k\sum_{j\geq 1}(i^\nu + k^\nu)c_j + \delta_{k,1} . \tag{4.82}$$

[8] We tacitly assume that periphery diffusion smooths the boundary so that the perimeter of the island scales as R.

[9] The term $(R_i + R_j)^{d-2}$ in the Smoluchowski formula should be replaced by the slowly varying factor $1/\ln(R_i + R_j)$ that we ignore in this presentation.

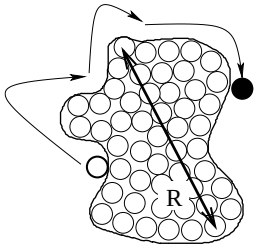


Figure 4.8: Left: Schematic illustration of effective island diffusion by the motion of an adatom at the periphery of an island.

We now introduce the two generating functions

$$\mathcal{A}(z) = \sum_{k\geq 1} k^\nu c_k\, z^k\,, \quad \mathcal{C}(z) = \sum_{k\geq 1} c_k\, z^k\,, \tag{4.83}$$

as well as $A = \mathcal{A}(z=1) = \sum_{k\geq 1} k^\nu c_k$, to reduce the master equation (4.82) to

$$\mathcal{A}(z)\mathcal{C}(z) - \mathcal{A}(z)N - \mathcal{C}(z)A + z = 0\,. \tag{4.84}$$

Since the mass distribution decays algebraically when $\nu = 0$, it is natural that this decay holds when $\nu \neq 0$; thus we assume that $c_k \to C\,k^{-\tau}$ when $k \gg 1$. To determine the decay exponent τ and the amplitude C we determine the singular behavior of the generating functions from this hypothesized mass distribution. From appendix???, the behavior $c_k \to C\,k^{-\tau}$ when $k \gg 1$ is equivalent to the algebraic behavior of the generating functions $\mathcal{C}(z)$ and $\mathcal{A}(z)$ in the $z \uparrow 1$ limit:

$$\begin{aligned} \mathcal{A}(z) &= A + C\Gamma(1-\tau+\nu)\,(1-z)^{\tau-\nu-1} + \ldots \\ \mathcal{C}(z) &= N + C\Gamma(1-\tau)\,(1-z)^{\tau-1} + \ldots \end{aligned} \tag{4.85}$$

Substituting these expansions for $\mathcal{A}$ and $\mathcal{C}$ and into (4.84) and matching the constant terms as $z \to 1$ yields $AN = 1$. Matching then the first correction terms in $(1-z)$ and using the reflection formula for the gamma function, $\Gamma(z)\Gamma(1-z) = \pi \mathrm{cosec}\,\pi z$, we obtain the decay exponent τ and the amplitude C

$$\tau = \frac{3+\nu}{2}\,, \qquad C = \sqrt{\frac{1-\nu^2}{4\pi}\,\cos\left(\frac{\pi\nu}{2}\right)}\,. \tag{4.86}$$

We many now estimate the crossover time by the same reasoning that led to (4.77):

$$t = \sum_{k=1}^{\infty} k c_k(t) \approx \sum_{k=1}^{k_*} k c_k \sim \sum_{k=1}^{k_*} k^{1-\tau} \sim k_*^{2-\tau} = k_*^{(1-\nu)/2}\,,$$

to give the characteristic mass

$$k_* \sim t^{2/(1-\nu)}\,. \tag{4.87}$$

Our analysis has relied on the assumption that the system reaches a stationary state. Specifically, we have tacitly assumed that both the sums $N = \sum c_k$ and $A = \sum k^\nu c_k$ converge, which means that the decay exponent should simultaneously satisfy $\tau > 1$ and $\tau > 1+\nu$. These inequalities in conjunction with $\tau = (3+\nu)/2$ impose the bounds on the mobility exponent

$$-1 < \nu < 1\,, \tag{4.88}$$

for the steady state to exist. When $\nu \to -\infty$, aggregates of any size are relatively immobile. In the extreme case of $\nu = -\infty$, *any* aggregate is immobile and this limit leads to the phenomenologically rich island growth model.

4.4 Island Growth Model

In the previous section, the mobility of an island was a rapidly decreasing function of its size. The *island growth* model represents the extreme limit in which islands of mass 2 or greater are immobile. As a result, islands grow only by the *addition* of mobile adatoms to their boundaries; there is no longer any aggregation *per se*. The island growth model has dramatically different behavior than that in the generalized sum kernel because a steady state no longer occurs. This model also nicely illustrates the power and the limitations of a scaling analysis for the mass distribution.

The elemental steps of island growth are:

$$0 \xrightarrow{F} A_1 \qquad A_1 + A_1 \to A_2 \qquad A_1 + A_k \to A_{k+1},$$

where F is the deposition rate. When only monomers are mobile, the reaction kernel is $K_{ij} \sim D_i + D_j \sim D(\delta_{i,1} + \delta_{j,1})$. This leads to the master equations

$$\begin{aligned} \frac{dc_1}{dt} &= -c_1^2 - c_1 \sum_{k=1}^{\infty} c_k + F \\ \frac{dc_k}{dt} &= c_1(c_{k-1} - c_k) \quad k \geq 2. \end{aligned} \tag{4.89}$$

Here we absorb the diffusivity D into the time variable and the parameter F becomes the ratio of the deposition rate (measured in number of adatoms per site per unit time) to the diffusivity D. In experimental applications the (dimensionless) parameter F is usually small and can be varied over a wide range. The master equations (4.89) involve several important assumptions that should be highlighted at the outset:

- Dilute islands. This limit corresponds to the submonolayer regime defined by $Ft \ll 1$ where islands are widely separated so that there is no possibility of proximal merging.
- Point-like (single-site) islands. This assumption is not as drastic as might appear at first sight because in two dimensions the reaction rate depends only logarithmically on the island radius.
- Mean-field master equations. This assumption is difficult to justify. Immobile islands can be viewed as "traps" and for the simplest example of randomly-distribution traps (see Chapter 9), fluctuation effects arise in all spatial dimensions. Hence the applicability of a mean-field description for two dimensions is questionable. Moreover, the input generates traps dynamically and its role on the applicability of the master equations is unknown.
- Freely diffusing adatoms and stable, immobile islands. This assumption is certainly questionable — small islands may be unstable, and even stable islands may undergo an effective diffusion.

To solve the master equations (4.89) it is most useful to first sum them to obtain the rate equation for the total density N:

$$\frac{dN}{dt} = F - c_1 N. \tag{4.90}$$

It is also helpful to partition the system into monomers and immobile islands — those of mass 2 or greater. Let I be the density of these immobile islands, $I = \sum_{k\geq 2} c_k$. Then monomers and islands satisfy the coupled equations

$$\begin{aligned} \frac{dc_1}{dt} &= -2c_1^2 - c_1 I + F \\ \frac{dI}{dt} &= c_1^2. \end{aligned} \tag{4.91}$$

One useful consequence of this partitioning is that it is obvious that the island density monotonically increases with time, and that $c_1 \to 0$ as $t \to \infty$. To verify that this latter assertion is true suppose the opposite; namely, $c_1 \to$ const. as $t \to \infty$. Then I would grow linearly with time, and the leading behavior of the monomer equation would be $\dot{c}_1 \approx -c_1 I$, which gives the contradiction that c_1 vanishes as $t \to \infty$. Thus

we conclude that $c_1 \to 0$ for $t \to \infty$. Using this result, the leading terms in the master equation for c_1 are $F - c_1 I = 0$, which gives $c_1 \simeq F/I$. Substituting this result into the second of (4.91) then gives

$$I(t) = (3t)^{1/3} F^{2/3}\,, \qquad c_1(t) \simeq \frac{F^{1/3}}{(3t)^{1/3}}. \tag{4.92}$$

These growth laws cannot hold indefinitely because the end of the submonolayer regime, in which there is less than adatom per adsorption site on average, is eventually reached. This regime obviously ends before $t_{\max} \sim F^{-1}$. Thus the maximal island density at the end of the submonolayer regime is

$$I_{\max} \sim F^{1/3}. \tag{4.93}$$

The theory that we now present applies only for $t < t_{\max}$.

To solve for the island densities we introduce the auxiliary time variable $\tau = \int_0^t c_1(t')\,dt'$ to recast the master equations (4.89) for immobile islands ($k \geq 2$) to those of the Poisson process

$$\frac{dc_k}{d\tau} = c_{k-1} - c_k\,. \tag{4.94}$$

To understand the asymptotic behavior, we treat k as continuous and approximate the difference by derivative to obtain the linear wave equation

$$\left(\frac{\partial}{\partial\tau} + \frac{\partial}{\partial k}\right) c_k(\tau) = 0. \tag{4.95}$$

The general solution is $c_k(\tau) = f(\tau - k)$, where f is an arbitrary function that is determined by matching the solution to c_1, which plays a role of a boundary condition. Since $c_1(\tau) = f(\tau - 1)$, the solution simply is $c_{k+1}(\tau) = c_1(\tau - k)$. Using the definition of τ and Eq. (4.92) for c_1, we have

$$\tau \simeq \frac{1}{2}\,F^{1/3}(3t)^{2/3}\,, \qquad c_1(\tau) \simeq F^{1/2}(2\tau)^{-1/2}\,, \tag{4.96}$$

which then gives

$$c_k(\tau) \simeq \frac{F^{1/2}}{\sqrt{2(\tau - k)}}\,. \tag{4.97}$$

The salient feature of this result is that the island distribution does not reach a steady state but rather is characterized by wave propagation.

While this wave solution works well over an intermediate range of island sizes k, it cannot apply at the extremes of k. For small k, the replacement of $c_k - c_{k-1}$ by $\frac{\partial c}{\partial k}$ is unjustified. More seriously, (4.97) must be erroneous when $k \geq \tau$. However, we can solve (4.94) exactly by introducing the Laplace transform $\widehat{c}_k(s) = \int_0^\infty c_k(\tau)\, e^{-s\tau}\, d\tau$ to recast this equation into $s\widehat{c}_k(s) = \widehat{c}_{k-1}(s) - \widehat{c}_k(s)$, from which we obtain

$$\widehat{c}_{k+1}(s) = (s+1)^{-1}\,\widehat{c}_k(s) = \ldots = (s+1)^{-k}\,\widehat{c}_1(s).$$

Because $\widehat{c}_{k+1}(s)$ is expressed as a product, its inverse Laplace transform is just the convolution of the inverse transforms of the two factors in the product:

$$c_{k+2}(\tau) = \frac{1}{k!}\int_0^\tau c_1(\tau - u)\, u^k\, e^{-u}\, du, \tag{4.98}$$

where $u^k\, e^{-u}/k!$ is the inverse Laplace transform of $(s+1)^{-k}$. We now use the Laplace method to show that this exact solution approaches the approximate wave equation solution (4.97) when $k, \tau \to \infty$, with $k/\tau < 1$. For large k, we expand the sharply-peaked factor $u^k e^{-u} = e^{k\ln u - u}$ about its maximum at $u^* = k$ and extend the integration to an infinite range to give

$$c_{k+2}(\tau) \simeq \frac{e^{k\ln k - k}}{k!}\int_{-\infty}^{\infty} c_1(\tau - k - \epsilon)\, e^{-\epsilon^2/2k}\, d\epsilon\,, \tag{4.99}$$

where $\epsilon = u - k$. Compute this Gaussian integral and using Stirling's approximation for the factorial $k! \simeq \sqrt{2\pi k}\,(k/e)^k$ then gives $c_{k+2}(\tau) \simeq c_1(\tau - k)$.

The exact solution (4.98) also allows us to resolve the apparent singularity of the approximate solution (4.97) near $k = \tau$. Laplace's method still applies and it tells us that the change of behavior occurs in an inner layer of width $\sqrt{\tau}$ around $k \approx \tau$. This suggests to introduce the 'layer' variable

$$k \to \infty, \quad \tau \to \infty, \quad y \equiv \frac{k-\tau}{\sqrt{2\tau}} = \text{finite} \tag{4.100}$$

We write $u = k - w\sqrt{2\tau}$ and proceed in the same way as in calculation (4.99):

$$\begin{aligned} c_{k+2}(\tau) &\approx \frac{e^{k\ln k - k}}{k!}\sqrt{2\tau}\int_{\xi}^{\infty} c_1[\sqrt{2\tau}\,(w-y)]\,e^{-w^2}\,dw \\ &\approx \frac{1}{\sqrt{2\pi\tau}}\sqrt{2\tau}\,\frac{F^{1/2}}{(2\tau)^{1/4}}\int_{y}^{\infty}\frac{e^{-w^2}}{\sqrt{2(w-y)}}\,dw \end{aligned}$$

In the second step we additionally used the asymptotic (4.96) of the monomer density. Therefore within the inner layer (4.100) the density has a scaling form

$$c_k(\tau) = \frac{F^{1/2}}{(4\pi\tau)^{1/4}}\,G(y) \tag{4.101}$$

with scaling function

$$G(y) = \int_{y}^{\infty}\frac{dw}{\sqrt{w-y}}\,e^{-w^2} \tag{4.102}$$

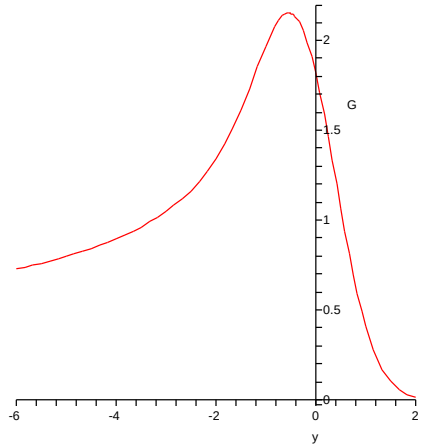


Figure 4.9: The scaled island size distribution $G(y)$ near the peak.

The singularity in the approximate solution is totally resolved — the maximal density does not diverge, it actually decays as $\tau^{-1/4} \propto t^{-1/6}$ although of course it greatly exceeds the density in the bulk of the mass distribution that decays as $t^{-1/3}$.

The island distribution near the peak (see Fig. 4.9) has the maximal density at $k - \tau \approx -0.541\sqrt{2\tau}$. When $y \to -\infty$, the right-hand side of Eq. (4.102) simplifies to

$$\int_{y}^{\infty}\frac{e^{-w^2}}{\sqrt{w-y}}\,dw \to \frac{1}{\sqrt{-y}}\int_{-\infty}^{\infty}e^{-w^2}\,dw = \sqrt{\frac{\pi}{-y}}$$

Using this in conjunction with (4.100), (4.101) we recover (4.97). Similarly when $y \to \infty$, the integral on the right-hand side of Eq. (4.102) is estimated by writing $w = y + x/(2y)$ to give

$$\int_{\xi}^{\infty}\frac{e^{-w^2}}{\sqrt{w-y}}\,dw = \frac{e^{-y^2}}{\sqrt{2y}}\int_{-\infty}^{\infty}e^{-x-x^2/(4y^2)}\,\frac{dx}{\sqrt{x}} = \sqrt{\frac{\pi}{2y}}\,e^{-y^2}$$

up to terms of the relative order $O(y^{-2})$ and smaller.

4.5 Finite Systems

Thus far, we've treated aggregation in the thermodynamic limit. However, real systems are finite. We now examine properties of aggregation that arise from finiteness. For example, aggregation eventually ends because all the mass condenses into a single cluster. What is the average condensation time? What is the distribution of the condensation time? How many clusters are present at time t and what is the distribution of the number of clusters? These types of questions become central in finite systems.

Finite-size aggregating systems are much harder to analyze than the corresponding infinite systems. This fact may seem puzzling at first sight because a system with just 2 or 3 clusters is certainly simpler than its infinite analog. However, the description of a system with, say, 77 clusters is quite challenging. The source of the difficulty stems from the fact that the full description of a finite aggregating system is encoded in the string of non-negative integers $(N_1, N_2, \ldots, N_N)$, where each $N_k(t)$ is the number of clusters of mass k at time t. These cluster densities must always satisfy the constraint $\sum_{k=1}^{N} kN_k = N$. In this section, we use N for the total initial number of monomers; we reserve the notation $N(t)$ for the cluster density in a infinite system.[10] For a finite system the quantities $N_k(t)$ are *random* so that we need to deal with the probability distribution for each $N_k(t)$. Moreover, the N_k's are coupled and one must study the probability distribution $P(N_1, N_2, \ldots, N_N; t)$. While it is straightforward to write equations that govern this probability distribution that happen to linear, an exact analysis of these multivariate equations is cumbersome even for the simplest reaction kernels.

For infinite systems, the quantities $N_k(t)$ are asymptotically *deterministic*. This determinism represents a huge simplification for the dynamics. There exist fluctuations, of course, but their relative magnitude vanishes in the $N \to \infty$ limit. Quantitatively, the anticipated behavior of $N_k(t)$ is

$$N_k(t) = Nc_k(t) + N^{1/2}\xi_k(t)\,,$$

with the densities $c_k(t)$ satisfying the deterministic master equations for the infinite-system limit. This decomposition of the densities into a deterministic part and much smaller fluctuating component is a natural approach for dealing with finite-size effects in non-equilibrium systems. In the following sections, however, we discuss examples where finiteness is an all-encompassing feature.

Constant kernel aggregation

The finite-size constant-kernel aggregating system is the only example that is readily soluble because one can ignore the masses of the clusters to address questions about the condensation time and other general features. The simplest quantity is the total number of clusters. The cluster number is a stochastic variable that changes from m to $m-1$ at a rate $r_m = m(m-1)/N$. here, the transition rate r_m must be proportional to the total number of distinct pairs, $\binom{m}{2}$, and we can determine the normalization by demanding that for $m = N = 2$ the rate should equal 1. Since the average time for the event $m \to m-1$ is $\Delta t_m = r_m^{-1}$, the time T_k until k clusters remain therefore is

$$\begin{aligned}
\langle t_k\rangle &= \sum_{m=N}^{k+1} \Delta t_m = N\left[\frac{1}{N(N-1)} + \frac{1}{(N-1)(N-2)} + \cdots + \frac{1}{k\cdot(k+1)}\right] \\
&= N\left[\left(\frac{1}{N-1} - \frac{1}{N}\right) + \left(\frac{1}{N-2} - \frac{1}{N-1}\right) + \cdots + \left(\frac{1}{k+1} - \frac{1}{k}\right)\right] \\
&= \frac{N}{k} - 1. \qquad (4.103)
\end{aligned}$$

Thus the average completion time until a single cluster remains is $T \equiv \langle t_1\rangle = N - 1$.

[10] As usual, we assume the monomer-only initial condition: $N_k(0) = N\delta_{k,1}$ so that the population therefore condenses into a single cluster of mass N at the end of the reaction.

At the next level of detail, we study the probability $P_m(t)$ that the system contains m clusters at time t. This probability evolves as

$$\frac{dP_m}{dt} = r_{m+1}P_{m+1} - r_m P_m, \tag{4.104}$$

subject to the initial condition $P_m(0) = \delta_{m,N}$. We impose the boundary conditions $P_{N+1} \equiv 0$, so that this equation applies for all $1 \le m \le N$. To solve (4.104), we Laplace transform it to give

$$(s + r_m)P_m(s) = \delta_{m,N} + r_{m+1}P_{m+1}(s). \tag{4.105}$$

For $m = N$ we get $P_N(s) = (r_N + s)^{-1}$, from which we then solve the rest of Eqs. (4.105) recursively. The last and most important quantity is $P_1(s)$, the Laplace transform of the probability that the system consists of a single cluster. This probability is given by

$$P_1(s) = \prod_{m=2}^{N} \frac{r_m}{s + r_m} = \prod_{m=2}^{N} \left[1 + \frac{sN}{m(m-1)}\right]^{-1}. \tag{4.106}$$

In the large-N limit, we may set to upper limit to infinity and use the identity

$$\prod_{m=2}^{\infty} \left[1 + \frac{x}{m(m-1)}\right]^{-1} = \pi x \, \sec\left(\frac{\pi}{2}\sqrt{1-4x}\right),$$

to express the solution in the compact form

$$P_1(s) \simeq \pi N s \ \sec\left(\frac{\pi}{2}\sqrt{1-4sN}\right). \tag{4.107}$$

The moments of the condensation time $\langle t_1^n \rangle$ may now be read off from the power-series expansion of the Laplace transform, $P_1(s) = \sum_{n\ge 0} \frac{(-s)^n}{n!}\langle t_1^n\rangle$. The leading behavior of these moments is $\langle t^n\rangle \simeq C_n N^n$, with the first three coefficients $C_1 = 1$, $C_2 = \frac{\pi^2}{3} - 2$, and $C_3 = 12 - \pi^2$. Because the Laplace transformed distribution (4.107) obeys scaling in the large-N limit, *i.e.*, $P_1(s) = \phi(z)$ with $z = sN$, the distribution of condensation times is a function of the scaled time tN^{-1},

$$P_1(t) = N^{-1}\Phi\left(tN^{-1}\right). \tag{4.108}$$

The limiting behavior of the scaling function $\Phi(x)$ can be obtained by inverting the Laplace transform and gives the asymptotic behavior

$$\Phi(x) \simeq \begin{cases} 6\, e^{-x} & x \to 0; \\ \frac{1}{4}\left(\frac{\pi}{x}\right)^{7/2}\, e^{-\pi^2/4x} & x \to \infty. \end{cases} \tag{4.109}$$

The standard $1/N$-expansion technique shows that the number distribution becomes Gaussian (see highlight)

$$P_m(t) \to \frac{1}{\sqrt{2\pi\Delta^2}} \exp\left[-\frac{(m - \langle m\rangle)^2}{2\Delta^2}\right], \tag{4.110}$$

in the thermodynamic limit, $N \to \infty$. The average number of clusters, $\langle m\rangle = Nc$ with $c = (1+t)^{-1}$, merely reproduces the infinite system size result. Fluctuations in the number of clusters, characterized by the variance $(\Delta m)^2 = N\sigma^2$ with $\sigma^2 = \frac{1}{3}[(1+t)^{-1} - (1+t)^{-4}]$, become negligible compared with the mean, $\Delta m/\langle m\rangle \sim N^{-1/2}$. Moreover, the number of clusters is a *self-averaging* quantity as m approaches the deterministic value $\langle m\rangle$ in the thermodynamic limit.

$1/N$-expansions

This canonical method for solving the master equation is demonstrated for (4.104). Since m is large, the continuum limit is taken (to leading order, $m^2 - m \cong m^2$):

$$\frac{d}{dt}P(m) = N^{-1}\left(\frac{\partial}{\partial m} + \frac{1}{2}\frac{\partial^2}{\partial m^2}\right)\left(m^2 P\right). \tag{4.111}$$

Anticipating that averages and variances grow linearly with N, a transformation of variables from the extensive (N-dependent) variable m to the intensive (N-independent) variable α is made

$$m = Nc + N^{1/2}\alpha. \tag{4.112}$$

We now seek the distribution $F(\alpha)$ for the initial condition $F_0(\alpha) = \delta(\alpha)$.
The original master equation (4.111) is transformed using $\partial/\partial m = N^{-1/2}\partial/\partial\alpha$ and $d/dt = \partial/\partial t - N^{1/2}(dc/dt)\partial/\partial\alpha$ as follows

$$\frac{\partial}{\partial t}F - N^{1/2}(dc/dt)\frac{\partial}{\partial\alpha}F = N^{-1}\left(N^{-1/2}\frac{\partial}{\partial\alpha} + \frac{1}{2}N^{-1}\frac{\partial^2}{\partial\alpha^2}\right)[(Nc + N^{1/2}\alpha)^2 F]. \tag{4.113}$$

The leading $\mathcal{O}(N^{1/2})$ order terms vanish because the concentration satisfies $dc/dt = -c^2$, while the next leading $\mathcal{O}(N^0)$ order terms vanish when the distribution $F(\alpha)$ satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t}F(\alpha) = 2c\frac{\partial}{\partial\alpha}\left[\alpha F(\alpha)\right] + \frac{1}{2}c^2\frac{\partial^2}{\partial\alpha^2}F(\alpha). \tag{4.114}$$

The solution to such a second order equation with linear coefficients in the first derivative term is always Gaussian (see van Kampen's book). Thus, it is characterized by the moments $\langle\alpha\rangle$ and $\sigma^2 \equiv \langle\alpha^2\rangle$.
Multiplying the Fokker-Planck equation by α and integrating by parts once yields $\frac{d}{dt}\langle\alpha\rangle = 0$. Since the average vanishes initially, $\langle\alpha(t)\rangle = 0$. Multiplying (4.114) by α^2 and integrating by parts twice, yields $\frac{d}{dt}\sigma^2 = -4c\,\sigma^2 + c^2$. Using $c = (1+t)^{-1}$ and the initial condition $\sigma^2(0) = 0$, fluctuations in the variable α are characterized by the variance $\sigma^2 = \frac{1}{3}[(1+t)^{-1} - (1+t)^{-4}]$. The variable α is Gaussian-distributed

$$F(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{\alpha^2}{2\sigma^2}\right]. \tag{4.115}$$

The essence of the transformation (4.112) is that it separates the stochastic part $N^{1/2}\alpha$ from the deterministic part Nc.

Product kernel aggregation

For an infinite system, a gel forms at time $t_g = 1$ that contains a finite fraction of the mass of the entire system. In contrast, for a finite system, the corresponding feature is the appearance of a "giant" cluster that is distinct from the rest of the mass distribution. What is the size of this giant cluster? When does it first appear? We can answer these questions in a simple way by exploiting results about the infinite system in conjunction with basic ideas from extreme-value statistics.

For a finite but large system, the expected number of clusters of mass k for $t \lesssim 1$ is, from the leading behavior of Eq. (4.41),

$$N_k \simeq N c_k(t) \simeq \frac{N}{\sqrt{2\pi}\,k^{5/2}}\,e^{-k(1-t)^2}.$$

There will then be a large number of monomers, dimer, trimers, *etc.*, but for sufficiently large k there will be few or even no clusters of this size. From basic extreme statistics considerations (see matters of technique), the size of the largest cluster $k_{\max}$ is determined by the criterion that there should be one cluster of the population whose size lies in the range $(k_{\max}, \infty)$. That is,

$$\sum_{k\geq k_{\max}} N_k \simeq \int_{k_{\max}}^{N} dk\,\frac{N}{\sqrt{2\pi}\,k^{5/2}}\,e^{-k(1-t)^2/2} = 1. \tag{4.116}$$

Prior to the gelation time, we may estimate this integral by noting that the exponential factor is dominant, so that all other factors in (4.116) can be evaluated at the lower limit. At the gelation time, one can evaluate the integral without any approximation. Beyond the gelation time, it is much simpler to use the fact that there exists a gel whose fraction g is determined by Eq. (4.43). Assembling the results of these small calculations, the mass of the largest cluster is asymptotically given by

$$k_{\max} \sim \begin{cases} (\ln N)/(1-t)^2 & t < 1 \\ N^{2/3} & t = 1 \\ Ng & t > 1. \end{cases} \tag{4.117}$$

For an infinite system, the gelation transition is sharp. On the other hand, a phase transition cannot occur in a finite system. Instead, a drastic change occurs within a narrow time range known as the *scaling window*; *finite-size scaling* refers to the study of critical behavior in this region. The width of the scaling window can be estimated by equating the expressions for $k_{\rm rmax}$ in (4.117) below and at $t = 1$ to give[11] $(1-t)^{-2} \sim N^{2/3}$. Consequently, the width of the scaling window is proportional to $N^{-1/3}$.

We can also estimate the condensation time from the exact expressions for the k-mer densities for the infinite system given in Eq. (4.38). These exact results tell us that most finite clusters are monomers in the long-time limit:

$$c_1 = e^{-t} \;\gg\; c_2 = \frac{1}{2}\, t\, e^{-2t} \;\gg\; c_3 = \frac{1}{6}\, t^2\, e^{-3t} \quad etc.$$

The condition that a single monomer remains in the system, $Nc_1 = 1$, provides the criterion that determines the condensation time to be

$$t_1 \simeq \ln N\,. \tag{4.118}$$

Condensation occurs much more quickly in product-kernel aggregation than constant-kernel aggregation because of the overwhelmingly large reactivity of the largest clusters. In contrast to constant-kernel aggregation, where the condensation time distribution (that is, it remains a random quantity in the $N \to \infty$ limit), for the product kernel aggregation the condensation time becomes deterministic in the $N \to \infty$ limit.

Instantaneous gelation

As mentioned in passing earlier in this chapter, there is exists the somewhat pathological, but nevertheless quite surprising, feature of instantaneous gelation for reaction kernels that obey the scaling $K(1,j) = K(j,1) \sim j^\nu$ with $\nu > 1$. That is, the gel appears at a time $t = 0^+$! The finite system provides a deeper understanding of this peculiar phenomenon.

For concreteness, we consider generalized product kernels of the form $K(i,j) = (ij)^\lambda$ with $\lambda > 1$. Then the master equations (4.1) for the infinite system become

$$\frac{dc_k}{dt} = \frac{1}{2} \sum_{i+j=k} (ij)^\lambda\, c_i\, c_j - k^\lambda\, c_k \sum_i l^\lambda c_i. \tag{4.119}$$

We assume that N is very large, so the master equations (with $c_k = N_k/N$, where N_k is the *average* number of clusters of mass k) should provide a good approximation. At short times, we may neglect the loss terms and we immediately find the densities grow with time as $c_k \simeq A_k t^{k-1}$ as in the case of the constant reaction kernel. The coefficients satisfy the recursion relations

$$(k-1)A_k = \frac{1}{2} \sum_{i+j=k} (ij)^\lambda\, A_i\, A_j \tag{4.120}$$

for $k \geq 2$ and with $A_1 = 1$. For large k, the dominant contribution to the sum is $(k-1)^\lambda A_{k-1}$. Keeping only this leading term, the recursion simplifies to $(k-1)A_k = (k-1)^\lambda A_{k-1}$, so that $A_k \sim [(k-1)!]^{\lambda-1}$. Then the average number of clusters of mass k is

$$N_k = Nc_k \sim N[(k-1)!]^{\lambda-1}\, t^{k-1}. \tag{4.121}$$

[11] The same result is obtained by equating the estimates (4.117) at and above $t = 1$, and using $g \sim t - 1$.

The first dimer appears when the condition $N_2 = 1$ holds, which gives the dimer appearance time $t_2 \sim N^{-1}$. Generally, the time of the first appearance of a cluster of size $k+1$ is given by $N_{k+1} = 1$, which gives the condition

$$t_{k+1} \sim (k!)^{-(\lambda-1)/k} N^{-1/k}. \tag{4.122}$$

Physically, the time for the first appearance of a k-mer should increase monotonically with k. However, (4.122) is monotonic only for sufficiently small k. This is a manifestation of the break-down of the master equations. Statistical fluctuations in the number of clusters of a given size are overwhelming, and average quantities such as $\langle N_k \rangle$ do not faithfully characterize the behavior (instead, the detailed probability distribution is needed). We anticipate that the largest cluster at this time, the giant cluster, absorbs the entire system mass.

The gelation time t_g can be estimated from the condition $t_{k_g} = t_{k_g+1}$. Using the Stirling formula, the size of the giant cluster grows logarithmically with the system size,

$$k_g \sim \ln N. \tag{4.123}$$

The giant cluster nucleates at a size that is much smaller compare with the product kernel case. Moreover, the gelation time $T_g \equiv t_{k_g}$ is $T_g \sim [\ln N]^{-(\lambda-1)}$. Once nucleated, the giant cluster grows according to $\frac{dk}{dt} = N^{-1}(N-k)k^\lambda \sim k^\lambda$, and the time for this growth process is found by integration,

$$T_N - T_g \sim \int_{k_g}^{N} \frac{dt}{dk} \sim \int_{k_g}^{N} dk\, k^{-\lambda} \sim k_g^{1-\lambda} - N^{1-\lambda} \sim k_g^{1-\lambda}. \tag{4.124}$$

Therefore, the condensation time is of the same order as the gelation time, $T_N \sim T_g$, so it vanishes logarithmically with system size,

$$T_N \sim [\ln N]^{-(\lambda-1)}. \tag{4.125}$$

This extremely slow decay indicates that in practice, it may be difficult to distinguish instant gelation from ordinary gelation.

Problems

Section 4.1

1. Use the generating function technique to solve for the n^{th} term in the Fibonacci sequence, F_n. This sequence is defined by the recursion $F_n = F_{n-1} + F_{n-2}$, with the boundary conditions $F_1 = F_2 = 1$.

2. Investigate the initial condition $c_k(0) = 2^{-k}$ in constant-kernel aggregation. Solve for the cluster concentrations.

3. Determine the 3^{rd} and 4^{th} moments of the cluster mass distribution for product kernel aggregation with the monomer-only initial condition. Show that

$$M_3(t) = \begin{cases} (1-t)^{-3} & \text{for} \quad t < 1; \\ e^{2gt}(e^{gt} - t)^{-3} & \text{for} \quad t > 1; \end{cases}$$

and

$$M_4(t) = \begin{cases} (1+2t)(1-t)^{-5} & \text{for} \quad t < 1; \\ (e^{4gt} + 2te^{3gt})(e^{gt} - t)^{-5} & \text{for} \quad t > 1. \end{cases}$$

More generally, show that near the gel point the moments M_n diverge according to

$$M_n \simeq \frac{2^{n-2}\,\Gamma(n-3/2)}{\Gamma(1/2)}\,|1-t|^{-(2n-3)} \tag{4.126}$$

which follows from Eq. (4.42).

4. Consider the addition process that can be represented as $A_1 + A_k \to A_{k+1}$, where the rate of each reaction is the same, except for the reaction between monomers, $A_1 + A_1 \to A_2$ which proceeds at a twice larger rate. (Why?) This process represents the growth of immobile islands on a surface due to the irreversible attachment of mobile monomers (Section 4.4). Let the process begins with monomer-only initial condition.

   (a) Introduce an auxiliary time that allows to linearize the master equations.

   (b) Determine the island size distribution at infinite time.

   (c) Show that the approach to the final state is exponential.

5. Consider the same addition process as in the previous problem but assume that the reaction is $K_{ij} = i\delta_{j,1} + j\delta_{i,1}$. Show that

$$c_k(t) = \left[(1-e^{-t})^{k-1} - k^{-1}(1-e^{-t})^k\right](2-e^{-t})^{-k}$$

Section 4.2

6. Use the scaling approach to determine the mass distribution for constant-kernel aggregation; that is, solve Eq. (4.66) for the scaling function.

Section 4.3

In all problems for this section, the system is initially empty, the input is a time-independent source of monomers that has started at $t = 0$, the strength of the input is set to unity so that mass density is $M = t$.

7. Use a direct calculation to find the first few k-mer densities for constant kernel aggregation with input.

   (a) Show that the monomer density is given by

$$c_1 = \frac{1}{2}\left[\frac{t}{\cosh^2 t} + \tanh t\right].$$

   (b) Solve recursively Eqs. (4.71) and derive a formal solution

$$c_k(t) = \frac{1}{\cosh^2 t}\int_0^t dt'\,\cosh^2 t' \sum_{i+j=k} c_i(t')c_j(t').$$

8. Consider the sum kernel aggregation with input.

   (a) Verify that the total density is given by

$$N(t) = \int_0^t dt'\, e^{(t'^2-t^2)/2}$$

   Show that $N(t)$ exhibits a non-monotonous behavior — it grows, reaches a maximum, and then decays to zero, $N \simeq t^{-1}$ when $t \gg 1$.

   (b) Show that the density of monomers also initially increases, then decreases, and asymptotically decays as $c_1 \simeq t^{-1}$.

   (c) Show that the density of dimers decays as $c_2 \simeq t^{-3}$.

   (d) Verify that generally $c_k \simeq A_k\, t^{-(2k-1)}$ in the long time limit.

   (e) Find a recursion for the amplitudes A_k and show that these amplitudes form a sequence **A088716** from *The On-Line Encyclopedia of Integer Sequences*.

9. Investigate the product kernel aggregation with input.

   (a) Show that in the pre-gel regime, the second moment is $M_2 = \tan t$. Use this solution to argue that gelation occurs at $t_g = \pi/2$.

   (b) Show that $N = t - t^3/6$ in the pre-gel regime.

   (c) Show that $M_3 = \tan t + \frac{2}{3}\tan^3 t$ in the pre-gel regime.

   (d) Show that the density of monomers is

   $$c_1(t) = \int_0^t dt'\, e^{(t'^2 - t^2)/2}$$

   throughout the evolution.

Section 4.4

10. Verify that the scaling function (4.102) can be expressed via the modified Bessel function

$$G(y) = \sqrt{\frac{y}{2}}\, e^{-y^2/2}\, K_{1/4}(y^2/2)$$

11. Consider the model (4.129) with unstable dimers.

   (a) Show that the exact solution (4.127) has a scaling form

   $$c_k(\tau) = \frac{F^{2/3}\lambda^{1/3}}{(18\tau)^{1/3}}\, G(y)$$

   in the scaling region (4.100) with scaling function

   $$G(y) = \frac{1}{\sqrt{\pi}} \int_y^\infty \frac{dw}{(w-y)^{2/3}}\, e^{-w^2}$$

   (b) Plot $G(y)$.

   (c) Verify that the density c_k is maximal at $k - \tau \approx -0.333\sqrt{2\tau}$, and its value decays as $\tau^{-1/3} \propto t^{-1/4}$.

12. Assume that islands of mass $\geq n$ are stable, while lighter islands are unstable.

   (a) Show that $c_1 \propto t^{-1/(n+1)}$.

   (b) Show that $c_n \propto t^{-(n-1)/(n+1)}$.

   (c) Show that $I_{\max} \propto F^{(n-1)/(n+1)}$.

13. Investigate the effect of island instability in the island growth model Consider the simplest situation where dimers are unstable, namely each dimer can break into two mobile adatoms with rate λ, while all larger islands are stable. Proceed in the same way as for the case where dimers are stable and show that asymptotically $I = F^{3/4}(4t/\lambda)^{1/4}$, while $c_1 = F^{1/4}(4t/\lambda)^{-1/4}$ and $c_2 = F^{1/2}(4t\lambda)^{-1/2}$. Also show that the exact solution for the densities of stable islands ($k \geq 3$) is

$$c_{k+3}(\tau) = \frac{1}{k!} \int_0^\tau du\, c_2(\tau - u) u^k e^{-u}\ . \tag{4.127}$$

Finally show that the maximal island density which is reached at the end of the submonolayer regime now follows the scaling law

$$I_{\max} \propto F^{1/2}\ . \tag{4.128}$$

Solution: The rate equations for monomers and dimers are

$$\begin{aligned} \frac{dc_1}{dt} &= F - c_1(c_1 + N) + 2\lambda c_2 \\ \frac{dc_k}{dt} &= c_1(c_1 - c_2) - \lambda c_2 \end{aligned} \tag{4.129}$$

while the densities of stable islands continue to evolve according to Eqs. (4.89). The total density of stable islands obeys

$$\frac{dI}{dt} = c_1 c_2 \,, \qquad I \equiv \sum_{k\geq 3} c_k = N - c_1 - c_2 \,. \tag{4.130}$$

The leading behavior of these equations gives that $I(t)$ grows while c_1, c_2 decay, so from (4.129) we deduce the asymptotic relations $c_1 = F/I$ and $\lambda c_2 = c_1^2$. Substituting these results into (4.130) gives a differential equation for $I(t)$ whose solution is

$$I = F^{3/4}(4t/\lambda)^{1/4} \,, \tag{4.131}$$

while the asymptotic densities of adatoms and dimers are

$$c_1 = F^{1/4}(4t/\lambda)^{-1/4} \,, \quad c_2 = F^{1/2}(4t\lambda)^{-1/2} \,. \tag{4.132}$$

Following the analysis method given in Sec. 4.4, the approximate wave equation solution is now $c_{k+2}(\tau) = c_2(\tau - k)$, where

$$\tau = \frac{4}{3}\left(\frac{\lambda F}{4}\right)^{1/4} t^{3/4} \,, \quad c_2(\tau) = F^{2/3}\lambda^{1/3}(3\tau)^{-2/3} \,,$$

while the exact solution for all stable island densities ($k \geq 3$) is

$$c_{k+3}(\tau) = \frac{1}{k!}\int_0^\tau du\, c_2(\tau - u) u^k e^{-u} \tag{4.133}$$

Note that in addition to different time dependence, the maximal island density which is reached at the end of the submonolayer regime now follows a different scaling law

$$I_{\max} \propto F^{1/2} \tag{4.134}$$

Section 4.5

14. Determine the mass of the largest cluster in a finite system in constant-kernel aggregation.

Chapter 5

FRAGMENTATION

In fragmentation, an object continuously breaks into an increasing number of smaller pieces by external driving. Fragmentation is ubiquitous in nature. At geological scales, fragmentation is responsible for sand grains on beaches and for boulder fields. At the molecular level, chemical bond breaking underlies polymer degradation and the consumption of material in combustion. A fragmentation event can be visualized as

$$A_x \xrightarrow{a(x)} A_{x_1} + A_{x_2} + \ldots + A_{x_n}\,,$$

in which an object of mass x breaks in a mass-conserving way at a rate $a(x)$ into n fragments of masses $x_1, x_2, \ldots, x_n$, with $B(x_i|x)$ the production rate of a daughter fragment of mass x_i from a fragment of mass x (Fig. 5.1). Here n may be fixed or varying in each breaking event. As in the complementary process of aggregation, we want to understand which microscopic features of the breakup rates $a(x)$ and $B(x_i|x)$ determine the distribution of fragment masses, $c(x,t)$.

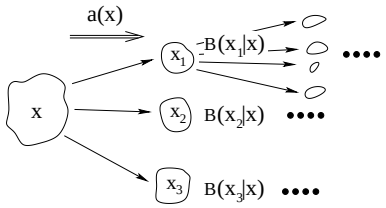


Figure 5.1: Schematic representation of fragmentation. A fragment of mass x breaks at an overall rate $a(x)$ into 3 daughter fragments of masses x_i, with respective production rates $B(x_i|x)$.

5.1 The Master Equation

The evolution of the mass distribution is described by the master equation

$$\frac{\partial c(x,t)}{\partial t} = -a(x)\,c(x,t) + \int_x^\infty c(y,t)\,a(y)\,B(x|y)\,dy\,. \tag{5.1}$$

The first term on the right accounts for the loss of fragments of mass x due to their breakup at overall rate $a(x)$. The second term accounts for the gain of fragments of mass x from the breakup of objects of mass larger than x. This master equation description involves a number of basic approximations, including:

- Homogeneity: fragment densities independent of spatial position. This is the mean-field assumption.
- Shape independence. Fragment shape plays no role in the evolution.
- Linearity. The breakup properties of a given cluster does not depend on the state of any other clusters.

Binary Breakup

Binary breakup refers to the specific case where two fragments are produced in each breaking event. The master equations may be written in a form that explicitly highlights this binary feature:

$$\frac{\partial c(x,t)}{\partial t} = -c(x,t)\int_0^x F(y,x-y)\,dy + 2\int_x^\infty c(y,t)\,F(x,y-x)\,dy\,. \tag{5.2}$$

Here $F(u,v)$ is the overall rate at which a fragment of mass $u+v$ breaks up into fragments of masses u and v. The factor 2 in the last term accounts for the fact that either one of the two daughter fragments from the breakup of a cluster of mass y may have mass x. Comparing with Eq. (5.1), we have $a(x) = \int_0^x F(y,x-y)\,dy$. Symmetry in the interchange of the two daughter fragments also implies that $B(x|y) = B(y-x|y)$. It is natural that the overall breakup rate does not depend on the masses of the two output fragments, in which case $F(y,x-y)$ is independent of y. Then the coefficient multiplying $c(x,t)$ in the first term on the right side of Eq. (5.2) — the overall breakup rate $a(x)$ — becomes $xf(x)$. Similarly, the coefficient of the second term is just $f(y)$. Perhaps the simplest example of binary breakup is *random scission*, where the breakup rate equals the fragment mass, $a(x) = x$ (equivalently, $f(x) = 1$). This choice describes the depolymerization of a linear polymer in which each chemical bond breaks at a fixed rate. Equivalently, random scission can be viewed as the cutting of a line (Fig. 5.2), with cuts occurring at a fixed rate per unit length. For random scission, the master equations reduce to

$$\frac{\partial c(x,t)}{\partial t} = -xc(x,t) + 2\int_x^\infty c(y,t)dy\,, \tag{5.3}$$

where we assume a scission rate of 1 per unit length.

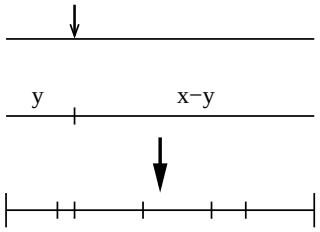


Figure 5.2: Random scission. An initial segment of length x is cut at a random point into two segments of lengths y and $x-y$. A series of such events can be viewed as the random deposition of cuts (bottom).

We now exploit this cutting picture to give a probabilistic derivation of the fragment mass distribution. Since cuts occur at random with rate 1, a segment of length Δx will be cut with probability $t\Delta x$ in a time t, and remains intact with probability $e^{-t\Delta x}$. There are then three types of fragments: (i) the initial segment remains intact, (ii) a fragment is at one end of the segment, (iii) a fragment is in the interior. We now evaluate the respective probabilities for these cases at time t, starting with an initial fragment of length L:

0 Intact segment. Since the average number of cuts is Lt, the segment remains intact with probability e^{-Lt}. The length distribution for such a segment is then $p_0(x) = e^{-Lt}\delta(x-L)$.

1 End fragment. An end fragment of length in the range $(x, x+dx)$ is created when there are no cuts within a distance x from the end (probability e^{-xt}), and the interval $(x, x+dx)$ contains a cut (probability $t\,dx$). The probability for this event is $p_1(x)dx = 2t\,e^{-xt}\,dx$, where the factor 2 accounts for the fragment being at either end. The length distribution for such segments is then $p_1(x) = 2te^{-xt}$.

2 Interior fragment. Such a fragment arises when there are two cuts in the intervals dx and dz at opposite ends of a otherwise intact fragment of length x. The probability for this event is $t^2\,dx\,dz\,e^{-xt}$. Integrating over all possible fragment positions ($x < z < L$) gives $p_2(x) = t^2(L-x)e^{-xt}$.

Combining these three events gives the probability for a fragment of length x at time t:

$$c(x,t) = e^{-xt}\left\{\delta(L-x) + \left[2t + t^2(L-x)\right]\Theta(L-x)\right\}. \tag{5.4}$$

It is easy to check that this expression is a solution of the master equation (5.3).

Because of the linearity of the master equations, we may average Eq. (5.4) over an arbitrary initial distribution $c(x,t=0) \equiv c_0(L)$ to give the general mass distribution

$$c(x,t) = e^{-xt}\left\{c_0(x) + \int_x^\infty c_0(y)\left[2t + t^2(y-x)\right]dy\right\}. \tag{5.5}$$

The controlling factor in this solution is the exponential e^{-xt} which leads to the distribution being non-zero only for $x < t^{-1}$. Thus the typical fragment mass shrinks and concomitantly the number of segments grows with time. The role of the initial boundaries of the segment thus become irrelevant as interior segments predominate the mass distribution. Thus keeping only the contribution of interior segments and using the fact that the typical segment length approaches zero as $1/t$, the asymptotic mass distribution reduces to

$$c(x,t) \simeq L_0\, t^2\, e^{-xt}, \tag{5.6}$$

where $L_0 = \int x\, c_0(x)\, dx$ is the average initial fragment mass. Notice that this asymptotic form can be written in the scaling form $c(x,t) \simeq L_0 s^{-2} f(x/s)$, with typical mass $s(t) = t^{-1}$ and scaling function $f(z) = e^{-z}$. For an exponential initial condition, $c_0(x) = e^{-xs}$, Eq. (5.5) gives the fragment mass distribution $c(x,t) = (1+t/s)^2\, e^{-x(t+s)}$. Thus an exponential distribution is preserved under fragmentation, but with a width that continuously decreases with time.

An alternative approach for determining the mass distribution is the *Charlesby Method*, which is essentially a Taylor expansion of the distribution in time. Even though a Taylor expansion ostensibly applies for short times, a key feature is that the series expansion coefficients can be obtained recursively. Consequently this method is exact, as nicely illustrated by the random scission model. As suggested by the form of the master equation, it is convenient to write $c(x,t)$ as $e^{-xt}F(x,t)$, a construction that leads to a master equation for F that contains only a gain term. We therefore write the $c(x,t)$ as the power series

$$c(x,t) = e^{-xt}\sum_{k=0}^{\infty} f_k(x)\, t^k. \tag{5.7}$$

The lowest-order term is dictated by the initial condition, $f_0(x) = c_0(x)$. Substituting this expansion into (5.3) and differentiating once with respect to x, yields the recursion relation for the expansion functions

$$\frac{d}{dx} f_k(x) = \frac{k-3}{k} f_{k-1}(x). \tag{5.8}$$

Integrating the first two of these equations, $f_1' = -2f_0$ and $f_2' = -\frac{1}{2}f_1$, yields

$$f_1(x) = 2\int_x^\infty c_0(x)dx, \qquad f_2(x) = \int_x^\infty (y-x)c_0(y)dy. \tag{5.9}$$

The terms with $k \geq 3$ all vanish and the resulting solution coincides with (5.5).

Finally, we present a direct solution to the master equation. The form of the first two terms in Eq. (5.3) leads to the integrating factor $e^{x(t+B)}$, with B a constant. This observation suggests that we seek a solution of the form $c(x,t) = A(t)\, e^{-x(t+B)}$. At this stage, it is convenient to differentiate the master equation with respect to x to convert it to the partial differential equation $c_{xt} = -xc_x - 3c$, where the subscripts denote partial differentiation. Substituting the ansatz $c(x,t) = A(t)\, e^{-x(t+B)}$ into this equation gives $\dot{A} = 2A/(t+B)$, with solution $A(t) = (t+B)^2$. The fragment mass distribution thus has the form $c(x,t) \propto (t+B)^2\, e^{-(t+B)x}$ and by matching with the initial condition, we reproduce the solution given in Eq. (5.5).

By this direct approach, we can also treat homogeneous scission, in which the overall breakup rate of a cluster is $a(x) = x^\lambda$ for general λ. The master equation is now

$$\frac{\partial c(x,t)}{\partial t} = -x^\lambda\, c(x,t) + 2\int_x^\infty y^{\lambda-1} c(y,t)\, dy. \tag{5.10}$$

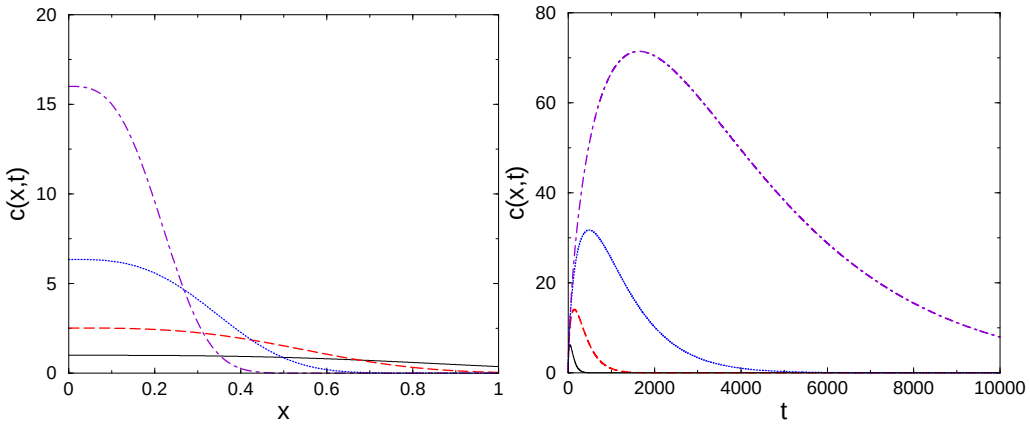


Figure 5.3: (Left) The asymptotic mass distribution (5.12) for the case $\lambda = 3$ as a function of x for $t = 1$ (solid), 4 (dash), 16 (dotted), and 64 (dot-dash). (Right) The same mass distribution as a function of t for $x = \frac{1}{4}$ (solid), $\frac{1}{6}$ (dash), $\frac{1}{9}$ (dotted), and $\frac{1}{13.5}$ (dot-dash).

We now seek a solution of the form $c(x,t) = A(t)e^{-x^\lambda(t+B)}$. Substituting this form into the master equation yields a soluble ordinary differential equation for $A(t)$, from which one obtains a special solution for $c(x,t)$,

$$c(x,t) = (1 + t/B_0)^{2/\lambda} e^{-(t+B_0)x^\lambda}. \tag{5.11}$$

By the linearity of the master equations, the general solution for $c(x,t)$ can be written as a linear combination, that is, an integral transform of Eq. (5.11), which, in turn, can be expressed in terms of special functions. In the long-time limit, the asymptotic behavior is

$$c(x,t) \propto t^{2/\lambda} \exp(-tx^\lambda). \tag{5.12}$$

A sketch of this mass distribution is shown in Fig. 5.3, both as a function of time for fixed mass, and as a function of mass at fixed times. Qualitatively, the behavior is in accord with simple-minded intuition. The population of a given (small) mass, x, initially grows due to the breakup of larger fragments. Eventually, however, the population at this mass decays when the production of x diminishes due to the depletion of larger fragments. This decay is asymptotically exponential, with the decay time varying inversely in the particle mass. For the distribution at fixed time, there is a steepening near the origin as a function of time, reflecting the eventual predominance by very small fragments. This small mass tail often has a power law form, as discussed previously. Notice also that this asymptotic form for $c(x,t)$ becomes pathological as $\lambda \to 0$. This is a signal of the shattering transition, which will be discussed in a later section.

A particularly interesting special case is when the homogeneity exponent equals zero. For $F(x,y) \propto (x+y)^{-1}$, the exact solution for the fragment mass distribution in the limit $\lambda \to 0$ is,

$$c(x,t) = e^{-t}\delta(x-l) + \frac{2te^{-t}}{l} \sum_{n=0}^{\infty} \frac{[2t\ln(l/x)]^n}{n!\,(n+1)!}. \tag{5.13}$$

In the small mass limit, a singularity in $c(x,t)$ develops due to the explosive growth in the number of very small fragments. The corresponding moments of this mass distribution are,

$$M_n(t) = l^n \exp[(1-n)t/(1+n)]. \tag{5.14}$$

Thus the total number of fragments, $M_0(t)$, grows exponentially in time, in contrast to the power law growth for $M_0(t)$ in the non-shattering regime. As we shall see, this solution corresponds to a system on the borderline between scaling and shattering.

5.2 Scaling

We now present the scaling approach to determine the asymptotic behavior of the fragment mass distribution. As in aggregation, the applicability of scaling rests on the hypothesis that a well-defined continuously-decreasing typical mass scale exists. Consequently, the mass distribution should be a function *only* of the ratio of the mass of a fragment to this typical mass. As in aggregation, we assume that the breakup rates are homogeneous functions: the overall breakup rate is $a(x) = x^\lambda$, thereby defining the homogeneity index λ, and the daughter rate $B(x|y)$ depends only on the ratio of the daughter mass to initial fragment mass, so that $B(x|y) \propto y^{-1}\, b(\frac{x}{y})$. By definition, the integral $\int_0^1 b(x)\, dx$ equals the average number of fragments produced in a single breakup event, and mass conservation imposes the condition $y = \int_0^y x\, B(x|y)\, dx$. In scaled form, this latter statement becomes $\int_0^1 x\, b(x)\, dx = 1$.

Let's start by investigating the moments of the mass distribution. We define the $\alpha^{\rm th}$ moment of this distribution, $M_\alpha(t) \equiv \int_0^\infty x^\alpha\, c(x,t)\, dx$, where the order α will typically be non-integer. To determine the time dependence of M_α, we multiply the master equation (5.1) by x^α and integrate over all x to give

$$\int_0^\infty \left[x^\alpha \frac{\partial c(x,t)}{\partial t} = -x^{\alpha+\lambda} c(x,t) + x^\alpha \int_x^\infty c(y,t) y^{\lambda-1} b(x/y)\, dy \right]. \tag{5.15}$$

We now apply a simple trick, that we will use repeatedly, to yield a closed equation for the moments. We simply interchange the integration order in the last term so that each integral involves a single variable (see Fig. 5.4). This interchange gives

$$\begin{aligned} \int_0^\infty x^\alpha\, dx \int_x^\infty c(y) y^{\lambda-1} b(x/y)\, dy &= \int_0^\infty y^{\lambda-1} c(y)\, dy \int_0^y x^\alpha b(x/y) dx \\ &= \int_0^\infty y^{\alpha+\lambda} c(y)\, dy \int_0^1 z^\alpha b(z)\, dz \qquad z \equiv x/y \\ &= M_{\alpha+\lambda} L_\alpha, \end{aligned}$$

where $L_\alpha \equiv \int_0^1 x^\alpha b(x)\, dx$ is the $\alpha^{\rm th}$ moment of the scaled daughter breakup rate.

Figure 5.4: Illustration of the interchange of integration order. Left: Integrating first over y from x to ∞ and then over all x. Right: Integrating first over x from 0 to y and then over all y.

The moments therefore evolve as

$$\dot{M}_\alpha = (L_\alpha - 1)\, M_{\alpha+\lambda}. \tag{5.16}$$

Starting with $M_1(t) = 1$, Eq. (5.16) gives $\dot{M}_{1-\lambda}(t) = (L_{1-\lambda} - 1)$, with solution $M_{1-\lambda}(t) = (L_{1-\lambda} - 1)t + c$, where $L_{1-\lambda} - 1 > 0$ for $\lambda < 1$ and c is another constant of order 1. Iterating leads to the asymptotic solution,

$$M_{1-k\lambda} \simeq \prod_{j=1}^{k} (L_{1-j\lambda} - 1)\, \frac{t^k}{k!} \;\propto\; t^k, \tag{5.17}$$

for the discrete set of index values $1 - k\lambda$. Thus the negative moments, $M_{1-k\lambda}$, with k a positive integer, are connected recursively and asymptotically grow as t^k. These negative moments play the role analogous to the positive integer moments of the mass distribution in aggregation.

We now apply scaling to determine the asymptotics of the fragment mass distribution itself. As in aggregation [see Eq. (4.64)], we write the scaling ansatz for the mass distribution in fragmentation as

$$c(x,t) = \frac{1}{s^2}\, f\Big(\frac{x}{s}\Big), \tag{5.18}$$

where $s(t)$ is the typical fragment mass, and the exponent -2 enforces mass conservation. As already highlighted in the corresponding discussion of the scaling theory of aggregation (Sec. 4.2), the basic feature of scaling for the mass distribution is it is a function of the ratio of the mass to the typical mass, rather than depending on mass and time separately. We now substitute the scaling ansatz (5.18) into the master equation (5.1) to give, after some simple algebra,

$$\frac{\dot{s}}{s^3}[f(u)+uf'(u)] = s^{\lambda-2}\left[-u^\lambda\, f(u) + \int_u^\infty f(v)\, v^{\lambda-1}\, b\Big(\frac{u}{v}\Big)\, dv\right].$$

Now it is a simple matter to separate out the time and scaled mass dependences to give

$$\begin{aligned} \dot{s}\, s^{-(1+\lambda)} &= -\omega \\ \omega[2f(u)+u\, f'(u)] &= -u^\lambda\, f(u) + \int_u^\infty f(v)\, v^{\lambda-1}\, b\Big(\frac{u}{v}\Big)\, dv, \end{aligned} \tag{5.19}$$

where ω is the separation constant that is positive so that the typical mass decreases with time.

From the first equation, the typical fragment mass has the asymptotic time dependence

$$s(t) \sim \begin{cases} (\lambda\omega t)^{-1/\lambda}, & \text{for } \lambda > 0, \\ e^{-\omega t}, & \text{for } \lambda = 0, \\ (t_c - t)^{1/|\lambda|}, & \text{for } \lambda < 0 \text{ and } t \to t_c, \end{cases} \tag{5.20}$$

where $t_c = [s(0)\lambda\omega]^{-1}$. Here is a nice demonstration of the utility of the scaling approach: with minimal knowledge about system details, we've determined the time dependence of the average fragment mass with little labor! The last case is particularly striking: when λ is less than zero, the fragmentation of the tiniest fragments occurs so quickly that the mean fragment mass vanishes at a finite time t_c. This singularity signals the *shattering* transition and is analogous to the gelation transition in aggregation with a homogeneity index larger than one. For $t > t_c$, a finite fraction of the mass is converted into a "dust" phase that consists of an infinite number of zero mass particles that manages to contain a finite fraction of the total mass.

Let's now determine the scaled fragment mass distribution. We start by converting the scaling equation for $f(u)$ [the second of Eqs. (5.19)] in the regime $\lambda > 0$ to a recursion for an infinite set of moments of the distribution, from which we then reconstruct the mass distribution. Thus we multiply both sides of Eq. (5.19) by u^α and integrate over all u. After an integration by parts, the left-hand side is simply $\omega(1-\alpha)m_\alpha$, where

$$m_\alpha \equiv \int_0^\infty x^\alpha\, f(x)\, dx \tag{5.21}$$

is the α^{th} moment of the scaling function. For the right-hand side, we interchange the integration order to transform (see again Fig. 5.4)

$$\int_0^\infty u^\alpha\, du \int_u^\infty v^{\lambda-1}\, f(v)\, b(u/v)\, dv \qquad \text{to} \qquad \int_0^\infty v^{\lambda-1}\, f(v)\, dv \int_0^v u^\alpha\, b(u/v)\, du, \tag{5.22}$$

and then express the second integral in terms of the moments of the scaled daughter breakup rate $L_\alpha = \int_0^1 x^\alpha b(x)\, dx$. We thereby obtain the moments of the scaling function recursively, in analogy with Eq. (5.16),

$$m_{\alpha+\lambda} = \omega\, \frac{\alpha-1}{1-L_\alpha}\, m_\alpha. \tag{5.23}$$

Comparing the definitions of the "bare" moments M_α and the scaled moments m_α, we see that they are related by $M_\alpha(t) = m_\alpha \times s(t)^{\alpha-1}$. The special case $\alpha = 0$ yields $N = M_0 = m_0/s$; that is, the average number of fragments N is inversely proportional to the typical mass.

To probe the large-mass tail of the mass distribution, one should focus on the high-order moments, that is, m_α for $\alpha \gg 1$. The most useful quantities are not the positive integer moments, but rather, from Eq. (5.23), the moments of order $\alpha = n\lambda$, with n a positive integer. For these α values, we iterate Eq. (5.23) and use the initial condition $m_0 = 1$ as well as the constraint $L_0 = 2$ for binary fragmentation, to give

$$m_\alpha = \omega^n \prod_{k=1}^{n-1} \frac{k\lambda - 1}{1 - L_{k\lambda}}. \tag{5.24}$$

For large n, the product is dominated by the factors with large k. In this limit, the form of $L_{k\lambda}$ is determined by the behavior of the breakup kernel $b(x)$ for $x \to 1$, namely, the limit where a fragment remains nearly intact in a breaking event. For this limit, a generic form of the kernel is $b(x) = b + \mathcal{O}((1-x)^\mu)$, for $x \to 1$, where $b \geq 0$ is the probability for no breaking to occur, and $\mu > 0$. Then $L_\alpha \simeq b/\alpha$ for large α. We now use the results of the highlight below to give

$$m_\alpha = \omega^n \lambda^{n-1} \frac{\Gamma(n-\frac{1}{\lambda})}{\Gamma(1-\frac{1}{\lambda})} \frac{\Gamma(n)\Gamma(1-\frac{b}{\lambda})}{\Gamma(n-\frac{b}{\lambda})} \ .$$

We now apply Eq. (4.3) to the ratio of gamma functions, as well as Stirling's approximation for $\alpha = n\lambda \to \infty$, to find after several straightforward steps,

$$m_\alpha \propto \left(\frac{\omega\alpha}{e}\right)^{\alpha/\lambda} \ \alpha^{[(b-1)/\lambda]-1/2} \ , \tag{5.25}$$

where a messy overall factor of order 1 has not been written.

Finite Products and Gamma Functions

We often encounter products of the form in Eq. (5.24) and are interested in this product for large n. What is the most direct way to obtain this asymptotics? One good way is to write the product in terms of gamma functions. For example, the numerator in (5.24) is

$$\prod_{k=1}^{n-1}(k\lambda - 1) = \lambda^{n-1}\prod_{k=1}^{n-1}(k-\tfrac{1}{\lambda}) = \lambda^{n-1}\frac{\Gamma(n-\frac{1}{\lambda})}{\Gamma(1-\frac{1}{\lambda})} \ .$$

An important step is to factor out a constant such that each successive term in the product differs by 1, which can then be expressed as a ratio of gamma functions. Similarly for the denominator in (5.24), we use $L_{\kappa\lambda} \sim b/k\lambda$ to give

$$\prod_{k=1}^{n-1}(1-L_{k\lambda}) = \prod_{k=1}^{n-1}(1-\tfrac{b}{k\lambda}) = \frac{1}{\Gamma(n)}\prod_{k=1}^{n-1}(k-\tfrac{b}{\lambda}) = \frac{\Gamma(n-\frac{b}{\lambda})}{\Gamma(n)\Gamma(1-\frac{b}{\lambda})} \ .$$

To extract the mass distribution from these moments, we work backwards and determine what distribution leads to the moments gives above. Because the moments involve the gamma functions, it is natural to anticipate that the scaled mass distribution has the form $f(x) = Ax^\mu \exp(-ax^\nu)$, where a, μ, and ν are constants, and normalization gives $A = \nu\, a^{(1+\mu)/\nu}/\Gamma(\frac{\mu+1}{\nu})$. For this distribution, the moments are then

$$m_\alpha = a^{-\alpha/\nu}\ \frac{\Gamma(\frac{\alpha+\mu+1}{\nu})}{\Gamma(\frac{\mu+1}{\lambda})} \longrightarrow \left(\frac{\omega\alpha}{e}\right)^{\alpha/\nu}\ \alpha^{[(\mu+1)/\nu]-1/2} \qquad \alpha \to \infty, \tag{5.26}$$

when we choose $a = 1/(\lambda\omega)$, $\mu = b-2$, and $\nu = \lambda$. The latter form then matches Eq. (5.25). Thus we infer that the scaled mass distribution has the asymptotic behavior

$$f(x) \sim x^{b-2}\ \exp[-x^\lambda/(\lambda\omega)] \qquad x \to \infty, \tag{5.27}$$

from which the fragment mass distribution is

$$c(x,t) \propto \frac{x^{b-2}}{s^b}\, e^{-t\,x^\lambda} \ . \tag{5.28}$$

This asymptotics matches the exact behavior from homogeneous binary fragmentation.

To determine the small-mass tail of the mass distribution, we now need the moments with large negative order. Accordingly, we choose $\alpha = 1 - n\lambda$ in Eq. (5.23) and iterate to arrive at the counterpart of Eq. (5.24),

$$m_{1-n\lambda} = \omega^{-n} \prod_{k=1}^{n} \frac{(L_{1-k\lambda}-1)}{k\lambda} \ . \tag{5.29}$$

Now the n dependence of $m_{1-n\lambda}$ for large n is now determined by the limiting form of $b(x)$ for x near 0, and there are two generic cases: breakup kernels with a strict lower cutoff and those without. A simple way to realize the former case, is to impose the condition that no clusters below a fixed relative mass $x_0 < 1$ are produced in a single breakup event. From the definition of L_α, the moment $L_{1-\alpha}$ then has the leading behavior $x_0^{-\alpha}/\alpha^{1+\mu}$ for large α. Substituting this expression into Eq. (5.29) yields

$$m_{1-n\lambda} \sim \frac{1}{(n!)^2\,(\omega\lambda^2)^n}\; x_0^{-\lambda-2\lambda-\cdots-n\lambda} \;\propto\; x_0^{-n^2\lambda/2} \sim e^{-\alpha^2 \ln x_0/2\lambda} \qquad \alpha\to\infty.$$

We again work backwards and determine the form of the mass distribution that corresponds to these moments. The quadratic dependence on α in the exponential suggests that the distribution is log-normal, and it is immediate to verify that indeed the asymptotic behavior of $f(x)$ is given by

$$f(x) \sim e^{-\lambda(\ln x)^2/(2\ln x_0)} \qquad\qquad x\to 0. \tag{5.30}$$

This log-normal form is to be expected and it also follows from a classic argument that makes use of a correspondence between fragmentation and a random multiplicative process. Schematically the mass of a particular fragment evolves as $x_0 \to x_1 \to x_2 \to \cdots \to x_N$, where the successive reduction factor, $r_k = x_k/x_{k-1}$, is a random variable with a well-behaved distribution. By the central limit theorem, $\log x_N = \sum_{k=0}^{N} \log r_k$ will be normally distributed, so that x_N will be distributed log-normally.

A second general class of mass distributions arises for breakup rates in which arbitrarily small clusters can be produced in a single event. This behavior is typified by the power law daughter breakup rate $b(x) \sim x^\nu$ for small x. From Eq. (5.23), m_α diverges whenever L_α diverges, which occurs for $\alpha < \alpha_c < 0$, since m_0 is finite. For α close to α_c the leading term in Eq. (5.23) dominates to give $m_\alpha \propto L_\alpha$. By comparing the definitions $m_\alpha = \int x^\alpha f(x)\,dx$ and $L_\alpha = \int x^\alpha b(x)\,dx$, it follows that $f(x)$ coincides with $b(x)$. Therefore if the daughter breakup rate has a power law behavior for $x \to 0$, then $f(x)$ has the *same* power-law tail.

5.3 Fragmentation with Input

Fragmentation with steady material input arises in many industrial processes, such as the crushing of a steady stream of raw mineral ore. As in the complementary problem of aggregation with input, we want to understand how input influences the fragment mass distribution. Let the input rate of objects of mass x be $I(x)$; for convenience we set the total input rate to one, $\int x\,I(x)\,dx = 1$. We analyze the effect of input for the simple case of random scission with overall breakup rate $a(x) = x$. The master equation is now

$$\frac{\partial c(x,t)}{\partial t} = -xc(x,t) + 2\int_x^\infty c(y,t)dy + I(x). \tag{5.31}$$

It is again convenient to analyze this master equation by studying the evolution of the moments of the mass distribution. Here we we study not the moments *per se*, but rather, the *Mellin transform* of the mass distribution. The Mellin transform is the same as the moment except for a shift in its order (see the highlight about the Mellin transform on the next page). The Mellin transform is often more convenient than the moments and we will use this transform in the rest of this chapter. Now using the same interchange of integration order illustrated in Fig. 5.4, the Mellin transform $c(s,t) = \int c(x,t)\,x^{s-1}dx$ evolves according to

$$\frac{\partial c(s,t)}{\partial t} = \left(\frac{2-s}{s}\right) c(s+1,t) + I(s), \tag{5.32}$$

where $I(s) = \int I(x)x^{s-1}dx$, is the Mellin transform of the input rate. The Mellin transform is well suited to deal with integral equations of the form (5.31) that typifies fragmentation because it converts the integral equation into a recursion formula that is often soluble. A similar philosophy was the basis for the generating function approach in aggregation, where a convolution is converted to a product by the generating function.

In the steady state, the Mellin transform in (5.31) satisfies (after shifting the index s by 1)

$$c(s) = \left(\frac{s-1}{s-3}\right) I(s-1). \tag{5.33}$$

Using basic properties of the Mellin transform given in the highlight below, the steady-state fragment mass distribution is

$$c(x) = x^{-1} I(x) + 2x^{-3} \int_x^\infty I(y) y \, dy. \tag{5.34}$$

In the limit of small fragment masses, the second term becomes dominant. Since the integral approaches 1 as $x \to 0$, the mass distribution has the universal algebraic tail

$$c(x) \simeq 2x^{-3} \ , \tag{5.35}$$

that is *independent* of the details of the input for well-behaved input functions.

This asymptotic behavior can also be obtained by the following heuristic argument. Because of the steady input, the total mass in the system, $M(t) = c(s = 2, t)$, grows linearly with time, $M(t) = t$. Similarly, from Eq. (5.31), the total number of fragments $N(t) = c(s = 1, t)$ satisfies $\dot{N}(t) = t + \mu$, where $\mu = \int I(x) dx$ is the number of fragments added per unit time. Consequently, $N(t) = \frac{1}{2}t^2 + \mu t$. The first two moments imply that the typical fragment mass is $M/N \sim t^{-1}$. Thus the mass distribution should approach the scaling form

$$c(x, t) \simeq t^3 F(xt) \qquad \text{as } t \to \infty, \tag{5.36}$$

with prefactor t^{-3} to ensure that the total mass in the system grows linearly with time. With this scaling form, a steady state is possible only when $F(z) \sim z^{-3}$ for small x. This fact implies that $c(x) \sim x^{-3}$.

The Mellin Transform

The Mellin transform $M(s)$ of a function $c(x)$ and its inverse are defined by

$$M(s) = \int_0^\infty c(x) \, x^{s-1} \, dx, \qquad c(x) = \int_{c-i\infty}^{c+i\infty} x^{-s} M(s) \, ds. \tag{5.37}$$

The Mellin transform is just the moment of order $s-1$ of the function $c(x)$ and appears when one is interested in the moments of the probability distribution for a positive definite quantity. The Mellin transform is also just a Laplace transform in disguise. By defining $x = e^{-y}$, Eq. (5.37) immediately becomes the Laplace transform $M(s) = \int_0^\infty c(y) \, e^{-sy} \, dy$. Thus we can adapt the known rules about inverting the Laplace transform to infer the inverse of Mellin transforms.

Here are some basic rules about the Mellin transform and its inverse:

1. If the moments m_α are finite for $\alpha < \alpha_c$ and infinite for $\alpha > \alpha_c$, then $c(x)$ asymptotically behaves as the power law, $c(x) \propto x^{-1-\alpha_c}$.
2. If $M(s)$ is the Mellin transform of $c(x)$, then $M(s-n)$ is the Mellin transform of $x^{-n} c(x)$. It is easily verified that this relation is obvious.
3. The Mellin transform of $g(x) = x^{-m} \int_x c(y) y^{m-1} dy$ is $\frac{M(s)}{s-m}$. This result relies on the interchange of integration order illustrated in Fig. 5.4. With this interchange, the Mellin transform of $g(x)$ is

$$\begin{aligned} \int_0^\infty x^{s-1} \, g(x) \, dx &= \int_0^\infty x^{s-1} \, x^{-m} \, dx \int_x^\infty y^{m-1} \, f(y) \, dy \\ &= \int_0^\infty y^{m-1} \, f(y) \, dy \int_0^\infty x^{s-m-1} \, dx \\ &= \frac{1}{s-m} \int_0^\infty y^{s-1} \, f(y) \, dy = \frac{1}{s-m} M(s) \end{aligned}$$

The full time-dependent solution can be obtained by the Charlesby method. We start by expanding the Mellin transform as a power series in time

$$c(s, t) = \sum_{k=1}^\infty \frac{t^k}{k!} \, c_k(s) \, , \tag{5.38}$$

and then solve the expansion functions $c_k(s)$ iteratively. For an initially empty system, this expansion does not contain a constant term, while for a steady input, the Mellin transform of the input contains only a

time-independent term $I(s)$. We now substitute the expansion (5.38) into Eq. (5.32) and equate terms with the same powers of time. This yields $c_1(s) = I(s)$, and $c_{k+1}(s) = -\frac{s-2}{s} c_k(s+1)$ for $k \geq 2$. Solving this set of equations recursively gives

$$c_{k+1}(s) = (-1)^k \frac{(s-1)(s-2)}{(s+k-1)(s+k-2)} I(s+k).$$

To invert this Mellin transform, we re-write $c_{k+1}(s)$ as the partial fraction expansion

$$c_{k+1}(s) = (-1)^k \left[1 - \frac{k(k+1)}{s+k-1} + \frac{k(k-1)}{s+k-2}\right] I(s+k).$$

From Eq. (5.38), the mass distribution can be written as a power series

$$c(x,t) = \sum_{k=0}^{\infty} \frac{t^{k+1}(-x)^k}{(k+1)!} c_k(x), \tag{5.39}$$

where the inverse transform of $c_{k+1}(s)$ has been conveniently written as $(-x)^k c_k(x)$. The three terms in the above expression for $c_{k+1}(s)$ can be inverted using the rules outlined in the Mellin transform highlight on page 83. The final expression for $c_k(x)$ is

$$c_k(x) = I_1(x) + \frac{k(k+1)}{x} I_2(x) + \frac{k(k-1)}{x^2} I_3(x), \tag{5.40}$$

with

$$I_1(x) = I(x) \quad I_2(x) = \int_x^{\infty} I(y)\, dy \quad I_3(x) = \int_x^{\infty} y f(y)\, dy. \tag{5.41}$$

Summing the three terms separately gives the mass distribution

$$c(x,t) = \sum_{k=1}^{3} t^k I_k(x) F_k(xt), \tag{5.42}$$

with the scaling functions

$$\begin{aligned} F_1(z) &= z^{-1}(1 - e^{-z}), \\ F_2(z) &= e^{-z}, \\ F_3(z) &= z^{-3}\left[2 - (2 + 2z + z^2)e^{-z}\right]. \end{aligned} \tag{5.43}$$

The function $F_3(z)$ has been obtained from the power series $F_3(z) = \sum_{k\geq 0} \frac{(-z)^k}{k!(k+3)}$. Eqs. (5.42) and (5.43) give the full time-dependent solution for an *arbitrary* time-independent input $I(x)$. In the limit $x \to 0$ and $t \to \infty$, with the scaling variable $z = xt$ kept finite, the third term in the sum of Eq. (5.42) dominates, and the anticipated scaling behavior of (5.36) is recovered with $F(z) = F_3(z)$.

5.4 Rectangular Fragmentation

It is natural to think of fragmentation as a geometric process, especially when one drops a plate that shatters when it hits the floor. Here we discuss a fragmentation model that incorporates the effects of shape in a minimalist way — rectangular fragmentation. A striking feature of this rectangular fragmentation is that satisfies an infinite number of "hidden" conservation laws. As a consequence, basic features of the mass distribution *multifractal* scaling (see the highlight on page 87 for a brief introduction).

Rectangular fragmentation is defined by the following breaking rule: given a population of rectangles, one is chosen with probability proportional to its area. For the selected rectangle, a random point inside the rectangle is chosen and then a crack through this point — either horizontal or vertical with equal probability

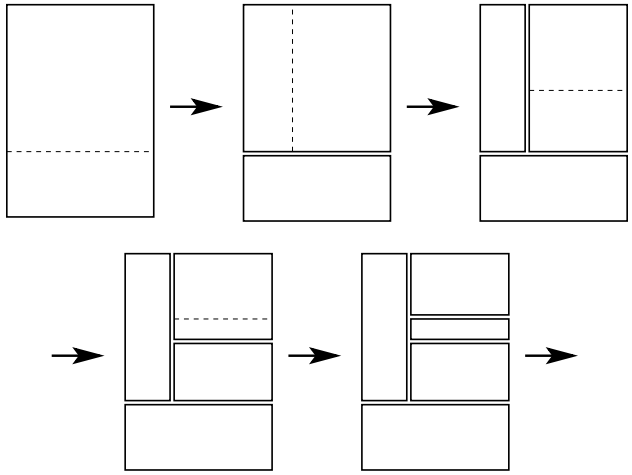


Figure 5.5: First few steps in rectangular fragmentation in two dimensions.

— fragments the rectangle into two smaller rectangles in an area-conserving manner (Fig. 5.5). Since one fragment is destroyed but two are created in each breaking event, the average number of fragments simply equals $N = 1 + t$.

The population of rectangles may be characterized by $c(x_1, x_2, t)$, the probability for a rectangle of length x_1 and width x_2 at time t. This distribution evolves in time according to the master equation

$$\frac{\partial c(x_1, x_2, t)}{\partial t} = -x_1 x_2\, c(x_1, x_2, t) + x_2 \int_{x_1}^{\infty} c(y_1, x_2, t)\, dy_1 + x_1 \int_{x_2}^{\infty} c(x_1, y_2, t)\, dy_2. \tag{5.44}$$

The first term accounts for the loss of a rectangle of area $A = x_1 x_2$ due to its fragmentation. Since the breaking point is chosen randomly inside from among the total area, the overall breaking rate is proportional to this area. The second term accounts for the gain of a fragment of length x_1 and height x_2 by vertically cracking a fragment of length $y_1 > x_1$ and height x_2. The prefactor x_2 accounts for the fact that the breaking point can be situated anywhere along the vertical crack. The last term accounts for the contribution due to a horizontal crack.

To solve (5.44), we introduce the 2-variable Mellin transform $M(s_1, s_2, t) \equiv \int\int x_1^{s_1-1} x_2^{s_2-1} c(x_1, x_2)\, dx_1\, dx_2$. Multiplying both sides of the master equation by $x_1^{s_1-1} x_2^{s_2-1}$ and integrating over all x_1 and x_2, the Mellin transform obeys the linear, non-local equation

$$\frac{\partial M(s_1, s_2, t)}{\partial t} = \left(\frac{1}{s_1} + \frac{1}{s_2} - 1\right) M(s_1 + 1, s_2 + 1, t). \tag{5.45}$$

From the definition of the Mellin transform, the moments of the fragment masses are then given by

$$\langle x_1^{n_1} x_2^{n_2} \rangle \equiv \frac{M(n_1 + 1, n_2 + 1)}{M(1, 1)}. \tag{5.46}$$

Thus, for example, the total area A of all fragments is just $M(2, 2)$, which is manifestly conserved in Eq. (5.45), while the average area is $\langle A \rangle = M(2, 2)/M(1, 1)$. Similarly, the total number of fragments $N = M(1, 1)$, satisfies $\frac{dN}{dt} = 1$, so that $N = (1 + t)$. Area conservation then implies that the average fragment area is $\langle A \rangle = \langle x_1 x_2 \rangle = (1 + t)^{-1}$.

A striking consequence of Eq. (5.45) is the existence of an infinity of conservation laws that are determined by the condition $(s_1^*)^{-1} + (s_2^*)^{-1} = 1$. These hidden conservation laws hold in an average sense. That is, while $M(s_1^*, s_2^*)$ does not strictly remain constant in each individual breaking event, $M(s_1^*, s_2^*)$, averaged over all realizations of rectangular fragmentation is conserved. However, the total area, $M(2, 2)$, is strictly conserved event by event.

Since the average mass decays algebraically with time in single-variable fragmentation with homogeneous breakup kernels, it is natural to assume that the Mellin transform also decays algebraically in time for rectangular fragmentation; that is $M(s_1, s_2) \sim t^{-\alpha(s_1,s_2)}$. Substituting this form into (5.45) gives $\alpha(s_1 + 1, s_2 + 1) = \alpha(s_1, s_2) + 1$. The exponent $\alpha(s_1, s_2)$ can now be determined by combining this recursion for the exponent with the conservation law $\alpha(s_1^*, s_2^*) = 0$ to give

$$\alpha(s_1^* + k, s_2^* + k) = k \qquad \text{for all} \qquad (s_1^*)^{-1} + (s_2^*)^{-1} = 1. \tag{5.47}$$

Thus the value of α at an arbitrary point (s_1, s_2) is just the horizontal (or vertical) distance from this point to the curve $(s_1)^{-1} + (s_2)^{-1} = 1$ (Fig. 5.6). This distance condition gives $(s_1 - \alpha) = s_1^*$ and $(s_2 - \alpha) = s_2^* = [1 - (s_1^*)^{-1}]^{-1}$. Eliminating s_1^* from these two equations, the exponent is the smaller root of the quadratic equation $(s_1 - \alpha)(s_2 - \alpha) - (s_1 - \alpha) - (s_2 - \alpha) = 0$, which gives

$$\alpha(s_1, s_2) = \frac{s_1 + s_2}{2} - 1 - \sqrt{\frac{(s_1 - s_2)^2}{4} + 1}\ .$$

Consequently, the asymptotic behavior of the moments is

$$\langle x_1^{s_1} x_2^{s_2} \rangle \equiv \frac{M(s_1 + 1, s_2 + 1)}{M(1, 1)} \sim t^{-\alpha(s_1+1,s_2+1)+\alpha(1,1)} \sim t^{-\alpha(s_1,s_2)-2}\ . \tag{5.48}$$

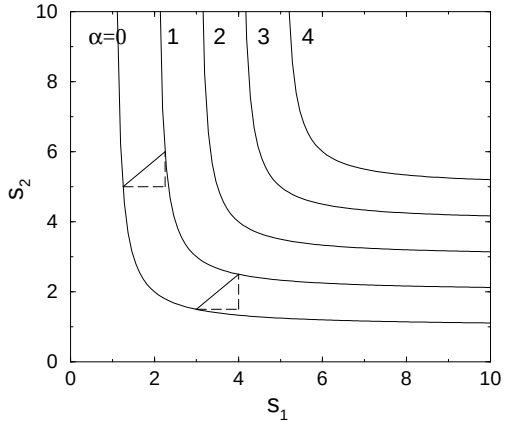


Figure 5.6: Loci of $\alpha(s_1, s_2) = 0$, 1, 2, 3, and 4 from Eq. (5.47). Each successive curve is shifted horizontally and vertically by 1 from its predecessor (dashed lines).

We now ask: what combinations of fragment length and width are conserved in rectangular fragmentation? The requirement that $M(s_1, s_2)$ is constant implies that $(s_1)^{-1} + (s_2)^{-1} = 1$. Therefore moments of the form

$$\int_0^\infty \int_0^\infty x_1^k\ x_2^{1/k}\, c(x_1, x_2)\, dx_1\, dx_2$$

are conserved for k. The case $k = 1$ gives the obvious conservation of the total area, but other k values lead to simple moment combinations whose physical meaning is not yet understood.

The non-linear spectrum of scaling exponents $\alpha(s_1, s_2)$ is an example of *multiscaling* or *multifractal scaling* in which high-order moments are not simply related to low-order moments, *viz.* $\langle x^n \rangle \gg \langle x \rangle^n$. A nice illustration of the consequences of this multiscaling arises in the moments of the fragment length:

$$\langle \ell^n \rangle \equiv \langle x_1^n \rangle \sim t^{-(n+2-\sqrt{n^2+4})/2}. \tag{5.49}$$

In particular, $\langle \ell \rangle \sim t^{-(3-\sqrt{5})/2} \sim t^{-.382}$ and $\langle \ell^2 \rangle^{1/2} \sim t^{-(2-\sqrt{2})/2} \sim t^{-.293}$. These moments decay slower than what one might naively anticipate from the behavior of the average area. From (5.48), the average area

decays as $\langle A\rangle = \langle x_1 x_2\rangle \sim t^{-1}$. Thus a natural length scale of the square-root of the area decays as $t^{-1/2}$. Another intriguing consequence of multiscaling is that a typical fragment becomes elongated (Fig. 5.5). We may quantify this visual asymmetry by the aspect ratio moments $\langle (x_1/x_2)^n\rangle$ which, from Eq. (5.48), diverge as $\langle (x_1/x_2)^n\rangle \sim t^{\sqrt{n^2+1}-1}$.

Scaling and Multiscaling

The notion of scaling, which permeates this book, typifies many complex systems. The notion of multiscaling is less well appreciated and it is worthwhile to give an example that illustrates the essential difference between scaling and multiscaling.

Consider the Gaussian probability distribution for a variable $x(t)$:

$$c(x,t) = \frac{1}{\sqrt{\pi t}}\, e^{-x^2/t}\ , \tag{5.50}$$

and let the exponent ν_n characterize the growth of the $n^{\rm th}$ moment of the variable x^2:

$$\langle x^{2n}\rangle = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi t}}\, x^{2n}\, e^{-x^2/t}\, dx = \frac{t^n}{\sqrt{4\pi}}\, \Gamma\left(n+\tfrac{1}{2}\right).$$

Thus the exponent $\nu_n = n$. This linear dependent of the exponent on the order of the moment defines scaling behavior. Another way to express this fact is that the ratios

$$\frac{\langle x^{2n}\rangle}{\langle x^2\rangle^n}$$

do not depend on time. Thus the $n^{\rm th}$ moment of x^2 scales as the $n^{\rm th}$ power of $\langle x^2\rangle$. Thus a single time scale characterizes all the moments. Now consider the same Gaussian distribution for the variable $x = \ln y$. Using the variable transformation $c(y) = c(x)\frac{dx}{dy}$, we have

$$c(y,t) = \frac{1}{\sqrt{\pi t}}\ \frac{1}{y}\, e^{-(\ln y)^2/t}. \tag{5.51}$$

For this distribution, the $n^{\rm th}$ moment is

$$\begin{aligned}
\langle y^n\rangle &= \frac{1}{\sqrt{\pi t}} \int_0^\infty e^{-(\ln y)^2/t + n\ln y}\, \frac{dy}{y} \\
&= \frac{1}{\sqrt{\pi t}} \int_0^\infty e^{-(\ln y/\sqrt{t} - n\sqrt{t}/2)^2 + n^2 t/4}\, d\ln y \\
&= \frac{1}{\sqrt{t}}\ e^{n^2 t/4}.
\end{aligned}$$

The ratio $\langle y^n\rangle/\langle y\rangle^n$ grows rapidly with time and thus a single scale is insufficient to characterize the distribution of the variable $\ln y$. This lack of scaling is termed *multiscaling* or *multifractal scaling*.

Although the length and aspect ratio exhibit multiscaling, the area distribution obeys conventional single-parameter scaling in which, again from Eq. (5.48), the area moments are characterized by a linear exponent spectrum: $\langle A^n\rangle \sim t^{-n}$. The area distribution is derived from the multivariate distribution using $c(A,t) = \int\int c(x_1,x_2)\,\delta(x_1x_2 - A)\,dx_1\,dx_2$. The area distribution obeys a scaling form as in Eq. (5.18): $c(x_1,x_2) \simeq t^{-2} f(At)$ with the scaling function

$$f(z) = 6\int_0^1 d\xi\,(\xi^{-1}-1)e^{-z/\xi}. \tag{5.52}$$

The area distribution diverges weakly: $f(z) \simeq 6\ln\frac{1}{z}$ at small areas, $z \ll 1$. In the opposite limit, there is an exponential decay, reminiscent of the one-dimensional case, but with an algebraic correction: $f(z) \simeq 6z^{-2}\exp(-z)$ as $z \gg 1$. The ordinary scaling behavior of the area reflects the fact that the fragmentation rate equals the area.

Problems

5.1 Exact Solutions

1. Solve the random scission model using the scaling ansatz $c(x,t) \simeq t^2\Phi(xt)$ (see matters of technique).
2. Obtain the leading asymptotic behavior of the moments for the random scission model.

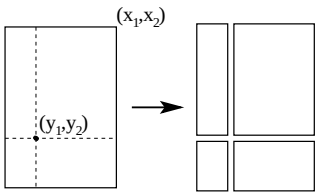
5.4 Rectangular Fragmentation

1. Determine the distribution of fragments of length x_1 and height x_2, $c(x_1,x_2)$, for the geometric fragmentation process illustrated below. Show that the two-variable Mellin transform $M(s_1,s_2)$ obeys

$$\frac{\partial M(s_1,s_2)}{\partial t} = \left(\frac{4}{s_1 s_2} - 1\right) M(s_1+1, s_2+1).$$

From this equation and using the assumption that $M(s_1,s_2) \sim t^{-\alpha(s_1,s_2)}$, show that

$$\alpha(n_1,n_2) = \left[(n_1+n_2) - \sqrt{(n_1-n_2)^2+16}\right]/2. \tag{5.53}$$



Chapter 6

ADSORPTION

This chapter is concerned with the kinetics of adsorption in which gas molecules impinge upon and then adsorb on a surface, or substrate. What is the rate at which adsorbed molecules fill the substrate? If the incident molecules are monomers that permanently attach to single adsorption sites on the surface and if there are no interactions between adsorbed monomers, then the fraction ρ of occupied sites increases at a rate proportional to the density of vacancies,

$$\frac{d\rho}{dt} = (1-\rho),$$

where we set the an intrinsic adsorption rate to 1 without loss of generality. The solution to this rate equation is $\rho(t) = 1-e^{-t}$, so that vacancies disappear exponentially in time. However, if each arriving molecule covers $k > 1$ substrate sites, then an unfilled region of less than k vacant sites can never be filled. The system therefore reaches a *jammed* state in which the substrate cannot accommodate additional adsorption, even though it is not completely filled. What is the filling fraction of this jammed state? What is the rate at which this final fraction is reached? These are the basic questions of adsorption kinetics.

6.1 Random Sequential Adsorption in One Dimension

Dimer adsorption

A simple example that exhibits non-trivial collective behavior is the irreversible and *random sequential adsorption* of dimers — molecules that occupy two adjacent sites of an infinite one-dimensional lattice (Fig. 6.1). We model the steady influx of molecules by adsorption attempts occurring one at a time at random locations on the substrate. An adsorption attempt is successful only if a dimer is incident onto two adjacent empty sites. If a dimer lands onto either two occupied sites or onto one occupied and one empty site, the attempt fails. That is, multilayer adsorption is forbidden, so that each site is either empty or contains 1 particle, but no more. The dimer coverage grows with time and eventually only isolated empty sites remain. When this occurs, the substrate is jammed and no further adsorption is possible.

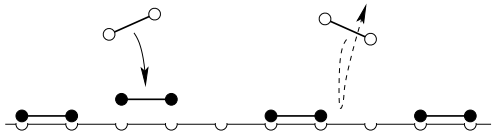


Figure 6.1: Irreversible dimer deposition. The dimer on the left successfully adsorbs onto two adjacent vacant substrate sites, while the dimer on the right does not adsorb.

For dimer adsorption, a jammed state consists of strings with an even number of occupied sites (•) that are separated by isolated vacant sites (◦), as illustrated below:

Figure 6.2: A jammed configuration in the random sequential adsorption of dimers.

In principle, the fraction of occupied sites in the jammed state, $\rho_{\rm jam} = \rho(t=\infty)$, can have any value between 2/3 and 1, with the two extreme limits achieved by the respective configurations:

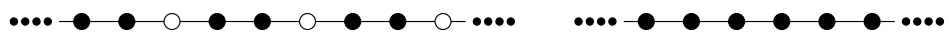


Figure 6.3: Minimum-density and maximum-density jammed dimer configurations.

A beautiful result, first derived by Flory, is that the value of $\rho_{\rm jam}$ in random sequential dimer adsorption is

$$\rho_{\rm jam} \equiv \rho(t=\infty) = 1 - e^{-2} = 0.864\,664\ldots \tag{6.1}$$

Flory's original derivation was based on enumerating all possible jammed configurations directly. Here we adopt a kinetic viewpoint and determine the time evolution of the coverage. The final coverage will then emerge as a direct consequence.

To determine the evolution of the substrate coverage, we need, in principle, the probabilities $P(\mathbf{n}, t)$ for the occupation state of each lattice site. Here $\mathbf{n} = \{n_j\}$, with $n_j = 1$ if the $j^{\rm th}$ site is occupied, and $n_j = 0$ if this site is empty. However these probabilities contain more information than necessary. What we really need are the *empty interval probabilities*, namely, the probability that a string of m consecutive sites are empty. We first define

$$E_m \equiv \text{prob}(\times \underbrace{\circ\cdots\circ}_{\rm m} \times) \equiv \mathcal{P}[\underbrace{\circ\cdots\circ}_{m}]$$

as the probability that there exists a string of m *consecutive* empty sites. Here the symbol $\times$ signifies that the state of the sites on the periphery of the m-interval are unspecified; they could be either occupied or empty. Consequently, E_m is also the probability to find an empty interval of length m *or greater*. In particular, E_1 is the density of empty sites and $\rho(t) = 1 - E_1$ is the density of occupied sites. Thus from the empty interval probabilities, we can obtain the particle density. This is one reason why the empty interval probabilities are so useful.

Figure 6.4: Changes in the empty interval probability E_m for $m = 4$. Shown are an adsorption event in the interior of the interval (left) and at the edge of the interval (right).

For irreversible dimer adsorption, the master equations that describe the evolution of the empty interval probabilities E_m are:

$$\frac{dE_m}{dt} = -(m-1)E_m - 2E_{m+1} \qquad m \geq 1. \tag{6.2}$$

The first term on the right side accounts for the loss of an m-interval due to the adsorption of dimers inside the interval. There are $m-1$ distinct locations at which the dimer can adsorb such that it lies entirely within the interval (Fig. 6.4). The second term accounts for the two adsorption events in which one end of the incident dimer is outside the m-interval. For these latter events, the empty interval must contain at least $m+1$ empty sites, hence the factor E_{m+1}. Notice that (6.2) contains only loss terms. This feature is a consequence of using the empty interval probabilities E_m as the basic dynamical variables. Pictorially, we are looking at the substrate through eyes that see only m consecutive sites at a time and E_m is merely the

fraction of these intervals that are empty. In this representation, there is no way to create an empty interval of length $\geq m$ by the adsorption of a dimer onto a still-larger empty interval.

It might seem more natural to write master equations for the *void densities*

$$V_m \equiv \mathcal{P}[\bullet \underbrace{\circ \cdots \circ}_{m} \bullet],$$

defined as the probability for m consecutive empty sites that are bounded on either site by an occupied site. The master equations that describe the evolution of the void densities V_m for irreversible dimer adsorption are:

$$\frac{dV_m}{dt} = -(m-1)V_m + 2\sum_{j=2}^{\infty} V_{m+j}. \tag{6.3}$$

The first term again accounts for the adsorption of a dimer in the interior of a void. The sum accounts for the *creation* of a void of m sites by the adsorbing a dimer into a void that contains $\geq m+2$ unoccupied sites. Notice the close correspondence between this master equation and those for fragmentation [*e.g.*, Eq. (5.3)]. There are no hard and fast rules for which set of quantities — E_m or V_m — are more useful for determining the dynamics of these type of kinetic problems. For adsorption, the master equations for E_m are typically easier to solve and we focus on these quantities in what follows:

Returning to the E_m, we now solve (6.2) for the initial condition of an initially empty system, $E_m(0) = 1$. Because an integrating factor for the master equations is $e^{-(m-1)t}$, this fact suggests seeking a solution of the form

$$E_m(t) = e^{-(m-1)t}\,\Phi(t), \tag{6.4}$$

where Φ coincides with $E_1(t)$, and with $\Phi(0) = 1$ to match the initial condition. Notice also that $E_1(t) = 1 - \rho(t)$, where $\rho(t)$ is the density of occupied sites. This connection allows us to determine how the surface coverage evolves. Using (6.4), the infinite set of master equations (6.2) miraculously reduces to the *single* equation $\frac{d\Phi}{dt} = -2e^{-t}\Phi$, whose solution immediately yields the empty interval probabilities

$$E_m(t) = e^{-(m-1)t-2\left(1-e^{-t}\right)}. \tag{6.5}$$

Empty gaps of length greater than 1 decay exponentially with time and only gaps of length 1 remain in the final jammed state. From (6.5), the density of such gaps is $E_1(\infty) = e^{-2}$, so that the jamming coverage is

$$\rho_{\rm jam} \equiv \rho(t=\infty) = 1 - e^{-2} = 0.864\,664\ldots \tag{6.6}$$

as first derived by Flory from a direct enumeration of all possible final state configurations.

While we have reproduced the classic Flory result with little labor, we also have much more — the coverage throughout the entire evolution:

$$\rho(t) = 1 - E_1(t) = 1 - e^{-2\left(1-e^{-t}\right)}. \tag{6.7}$$

The jamming coverage therefore approaches the jamming coverage exponentially in time, $\rho(\infty) - \rho(t) \to 2\,e^{-2}e^{-t}$, a feature that typifies lattice models of irreversible adsorption.

Adsorption of longer molecules

What happens if the incident molecules are k-mers that occupy k consecutive substrate sites? A nice illustration of the power of the master equation is that the coverage for this more general adsorption process can be obtained by a straightforward extension of the theory for dimers. The master equations for the empty interval probabilities E_m in k-mer adsorption are now:

$$\frac{dE_m}{dt} = \begin{cases} -(m-k+1)E_m(t) - 2\displaystyle\sum_{j=1}^{k-1} E_{m+j}(t) & m \geq k; \\[2ex] -(k-m+1)E_k(t) - 2\displaystyle\sum_{j=1}^{m-1} E_{k+j}(t) & m < k. \end{cases} \tag{6.8}$$

The terms in this equation mirror those in the master equation (6.2) for dimer adsorption. In the first line, the first term accounts for the $m-k+1$ distinct ways that a k-mer can adsorb in the interior of an m-site empty interval. The second term accounts for $2(k-1)$ ways that the k-mer can adsorb, with the k-mer partially outside and partially inside the original m-interval. For $m<k$, the first term accounts for the $k-m+1$ ways that the k-mer can cover the m-interval as well as $k-m$ sites outside the interval. The second term accounts for the ways in which the k-mer partially covers the interval. The equation for $m<k$ can be obtained quite simply by merely interchanging the roles of k and m in the equation for $m\geq k$.

For $m\geq k$, the structure of the equations again suggests the ansatz $E_m(t)=e^{-(m-k+1)t}\Phi(t)$ which reduces Eqs. (6.8) for $m\geq k$ to $\dot\Phi=-2\Phi\sum_{j=1}^{k-1}e^{-jt}$. Thus we obtain

$$E_m(t)=\exp\Big[-(m-k+1)t-2\sum_{j=1}^{k-1}\frac{1-e^{-jt}}{j}\Big]\qquad m\geq k. \tag{6.9}$$

To find the time dependence of the coverage, $\rho=1-E_1$ we use the fact that $E_1(t)$ satisfies the master equation $\dot E_1=-k\,E_k(t)$, with E_k determined from Eq. (6.9) with $m=k$. The coverage $\rho=1-E_1$, may then be expressed as

$$\rho(t)=k\int_0^t\exp\Big[-u-2\sum_{j=1}^{k-1}\frac{1-e^{-ju}}{j}\Big]\,du. \tag{6.10}$$

Numerical evaluation of this integral gives a jamming coverage that decreases monotonically with k (table 6.1 and Fig. 6.5). The jammed state becomes less full as k increases because empty regions as large as $k-1$ can arise. **seems incomplete**

k	$\rho_{\rm jam}$
1	1
2	0.864665
3	0.823653
4	0.803893
5	0.792276
∞	0.747597

Table 6.1: Jamming coverage for random sequential adsorption of k-mers in one dimension.

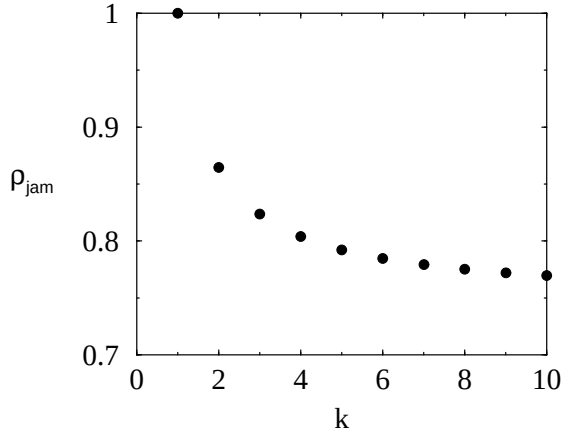


Figure 6.5: Jamming coverage for random sequential adsorption of k-mers in one dimension.

Irreversible car parking

The limit of k-mer adsorption with $k\to\infty$ defines the *car parking* problem. In this limit, the position of an adsorbed k-mer becomes continuous and it is simpler to think of unit-length "cars" that irreversibly

park anywhere along a one-dimensional curb (no marked parking spots) and then are abandoned. The only constraint is that cars cannot overlap; however a car *can* fit into a parking spot that is infinitesimally larger than the car itself. For this parking problem, the jamming coverage was found by Rényi to be $0.747597\ldots$, and is simply the $k\to\infty$ limit of the k-mer jamming coverage.

We may again solve this parking problem by the master equation approach. The length of a car is immaterial if we seek the fraction of the line that is covered, and for convenience we define the car length to be 1. The appropriate dynamical variable is $E(x,t)$, the probability that a randomly-chosen interval of length x is empty. As in the discrete case, this region may be part of an even longer empty interval. When the rate at which cars park equals 1, the master equation for $E(x,t)$ is [compare with Eq. (6.8)]

$$\frac{\partial E(x,t)}{\partial t} = \begin{cases} -(x-1)E(x,t) - 2\displaystyle\int_x^{x+1} E(y,t)\,dy & x>1 \\[2ex] -(1-x)E(1,t) - 2\displaystyle\int_1^{x+1} E(y,t)\,dy & x<1. \end{cases} \tag{6.11}$$

The terms in this master equation have direct counterparts with the terms in (6.8) for k-mer adsorption. For $x>1$, the first term on the right of Eq. (6.11) accounts for adsorption events that lie completely within the interval, while the second term accounts for adsorption events that partially overlap this interval. A similar correspondence also applies for the second line of (6.11).

To solve this master equation, consider first the regime $x>1$. As in the discrete case, we seek a solution of the form

$$E(x,t) = e^{-(x-1)t}\,E(1,t). \tag{6.12}$$

Substituting this expression into the first of Eqs. (6.11), the x-dependent terms cancel, and integrating the resulting equation for $E(1,t)$ gives

$$E(1,t) = \exp\left[-2\int_0^t \frac{1-e^{-u}}{u}\right]\,du,$$

with which Eq. (6.12) gives $E(x,t)$ for $x>1$. From the second of Eqs. (6.11), $\frac{\partial E(0,t)}{\partial t} = -E(1,t)$, from which the coverage $\rho(t) = 1-E(0,t)$ is

$$\rho(t) = \int_0^t \exp\left[-2\int_0^v du\,\frac{1-e^{-u}}{u}\right]\,dv. \tag{6.13}$$

For $t\to\infty$, numerical evaluation of this integral gives the jamming coverage $\rho(\infty) = 0.747597\ldots$.

A qualitative new feature of continuum car parking is that the approach to jamming is much slower than for adsorption of discrete molecules. Let's examine how $\rho(\infty)-\rho(t)$ vanishes as $t\to\infty$:

$$\begin{aligned} \rho(\infty)-\rho(t) &= \int_t^\infty \exp\left[-2\int_0^v \frac{du}{u}\,(1-e^{-u})\right]\,dv \\ &\sim \int_t^\infty \exp\left[-2\int_1^v \frac{du}{u}\right]\,dv \\ &\sim \int_t^\infty \frac{dv}{v^2} \propto \frac{1}{t}\ . \end{aligned}$$

The crucial step occurs in the second line. As discussed in Section A.2, we may replace the function $1-e^{-u}$, which gradually crosses over from 0 to 1 as u passes through 1, by a step cutoff that equals 0 for $u<1$ and equals 1 for $u>1$. Then the integral in the exponent is elementary and the asymptotic behavior follows straightforwardly. More precisely, the asymptotic behavior of $\int_0^t du\,\frac{1-e^{-u}}{u}$ is $\ln t + \gamma + t^{-1}e^{-t} + \ldots$, where $\gamma = 0.577215\ldots$ is Euler's constant. Thus

$$\rho(\infty) - \rho(t) \to t^{-1}\,e^{-2\gamma} \qquad t\to\infty;$$

the approach to jamming is much slower on continuous substrates than on discrete substrates.

6.2 Combinatorial Approach for Adsorption

Enumeration of dimer configurations

Flory originally determined the jamming coverage for dimer adsorption by a combinatorial method that we now present to contrast with the kinetic approach. In the combinatorial method, we write a recursion for the jamming density on a finite interval of length L in terms of jamming densities on shorter intervals. Let A_L be the average number of occupied sites in all final jammed states for dimer adsorption. If the first dimer lands on the sites $(i, i+1)$, then the remaining intervals of lengths $i-1$ and $L-i-1$ get filled independently. Therefore for $L \geq 2$, the coverage obeys the recurrence

$$\begin{aligned} A_L &= \frac{1}{L-1}\sum_{j=1}^{L-1}(A_{j-1} + 2 + A_{L-j-1}) \\ &= 2 + \frac{2}{L-1}\sum_{j=1}^{L-2} A_j, \end{aligned} \tag{6.14}$$

with $A_0 = A_1 = 0$. That is, the number of occupied sites in the jammed state equals 2 for the initial dimer plus the sum of the number of sites occupied in the two remaining subintervals. In the second line, we use the fact the two sums are identical. A crucial element of this recursion is the implicit assumption that dimers are added one at a time. Thus the final state will be the same as the jamming density that was obtained previously by the master equation description of the dimer deposition.

To solve for A_L we introduce the generating function $A(x) = \sum_{L=2}^{\infty} A_L x^{L-2}$, multiply Eq. (6.14) by $(L-1)\,x^{L-2}$, and sum over all L. The left-hand side is $\sum_2^{\infty}(L-1)\,x^{L-2}\,A_L = \frac{d(xA)}{dx}$. For the second term on the right-hand side, we interchange the order of summations (in close analogy with the interchange of integration order discussed on page 79) to give

$$\begin{aligned} 2\sum_{L=2}^{\infty} x^{L-2}\sum_{j=0}^{L-2} A_j &= 2\sum_{j=0}^{\infty} A_j \sum_{L=j+2}^{\infty} x^{L-2} \\ &= 2\sum_{j=0}^{\infty} A_j \frac{x^j}{1-x} \\ &= \frac{2x^2}{1-x}\,A(x), \end{aligned}$$

so that the recurrence (6.14) can now be recast as the differential equation

$$\frac{d}{dx}(xA) = \frac{2}{(1-x)^2} + \frac{2x^2}{1-x}\,A\,, \tag{6.15}$$

subject to the initial condition $A(x=0) = A_2 = 2$. The solution to Eq. (6.15) is

$$A(x) = \frac{1-e^{-2x}}{x(1-x)^2} = 2 + 2x + \frac{10}{3}x^2 + 4x^3 + \frac{74}{15}x^4 + \cdots\,, \tag{6.16}$$

and we can now read off the average number of occupied sites in the jammed state for small systems: $A_2 = 2$, $A_3 = 2$, $A_4 = \frac{10}{3}$, *etc.*

What is the meaning of these numbers? For example, consider $A_4 = \frac{10}{3}$, the average number of occupied sites on a 4-site chain. This value arises as follows: with probability $\frac{1}{3}$ the first dimer lands on the middle two sites and no further adsorption is possible; with probability $\frac{2}{3}$, the first dimer occupies one of the chain endpoints and the neighboring site. There are then still two more vacant adjacent sites that can accommodate one more dimer. Thus the average number of occupied sites is $\frac{1}{3}\times 2 + \frac{2}{3}\times 4 = \frac{10}{3}$. In contrast, in the equilibrium microcanonical ensemble, each jammed configuration occurs with the same probability, so that the average number of occupied sites equals 3.

As $L \to \infty$, we can obtain the series coefficients in the generating function by examining the behavior of $A(x)$ as $x \to 1$ from below. Now if $A_L \to L\rho$ as $L \to \infty$, with ρ a constant, then in this limit $A(x)$ would have the form

$$A(x) = \sum_{2}^{\infty} \rho L\, x^{L-2} = \frac{\rho}{x}\,\frac{d}{dx} \sum_{L=2}^{\infty} x^L$$
$$= \frac{\rho}{x(1-x)^2}(2x - x^2). \tag{6.17}$$

Comparing Eqs. (6.16) and (6.17) as $x \to 1$, we find $\rho = 1 - e^{-2} = 0.864\,664\ldots$, thus recovering the Flory result (6.1).

Phase space and broken ergodicity

A basic tenet of the microcanonical ensemble in equilibrium statistical physics is that of *equal a priori probability* — each microstate is realized with the same probability. This feature is termed *ergodicity* and it allows us to make the equivalence between the formal ensemble average of statistical mechanics and the time average in a real equilibrium system. Systems with jammed configurations do not fit into this picture as they do not uniformly sample the state space of all configurations. This phenomenon is known as *broken ergodicity*. Irreversible adsorption is an ideal setting to understand broken ergodicity because all states can be readily computed.

Let's start by counting the total number of jammed states in a finite system for irreversible dimer adsorption. Let F_L be the total number of jammed configurations on a finite chain of L sites. These jammed configurations can be divided into two categories: (i) those with the first site of the chain occupied and (ii) those with the first site empty. Configurations in these two categories obviously look like:

$$\bullet\bullet\underbrace{\times\cdots\times}_{L-2} \quad \text{and} \quad \circ\bullet\bullet\underbrace{\times\cdots\times}_{L-3},$$

respectively. Thus the first category consists of F_{L-2} configurations and the second category consists of F_{L-3} configurations. Thus F_L is determined by the recurrence

$$F_L = F_{L-2} + F_{L-3} \quad \text{for} \quad L > 2.$$

We write this Fibonacci-like recurrence in terms of the generating function and use the boundary conditions $F_0 = F_1 = F_2 = 1$ to give

$$F(x) = \sum_{L=0}^{\infty} F_L x^L = \frac{1 + x + x^2}{1 - x^2 - x^3}\ .$$

Formally, F_L is the L^{th} term in the generating function. For asymptotic behavior, it is sufficient to merely assume that $F_L \propto \zeta^L$ and substitute this assumption into the recursion relation for F_L to give $F_L = A\zeta^L + A_+\zeta_+^L + A_-\zeta_-^L$. Here ζ and $\zeta_\pm$ are the roots of the polynomial $z^3 - z - 1 = 0$ that explicitly are:

$$\zeta = a + b = 1.32472\ldots, \quad \zeta_\pm = a\, e^{\pm i\pi/3} + b\, e^{\pm 2i\pi/3}$$
$$\text{with} \quad a = \frac{1}{3}\left[\frac{27 + 3\sqrt{69}}{2}\right]^{1/3}, \quad b = \left[\frac{2}{27 + 3\sqrt{69}}\right]^{1/3}$$

Since $|\zeta_\pm| < 1$, the asymptotic behavior of the number of jammed configurations is given by

$$F_L \sim A\zeta^L, \tag{6.18}$$

where $A = (\zeta + \zeta^2 + \zeta^3)/(3 + 2\zeta) = 0.956614\ldots$ is obtained by matching the exact solution with the first few terms in the recursion relation. Thus the number of jammed configurations grows exponentially with the system size and there is an extensive packing entropy, $S = \ln F_L \sim L \ln \zeta$.

Next, we determine the number of configurations with a specified coverage. Let $F_{N,L}$ be the number of jammed configurations that contain N dimers in a system of size L. The number of dimers must be in the range $\lfloor (L+1)/3 \rfloor \le N \le \lfloor L/2 \rfloor$ for the configuration to actually be jammed, with $\lfloor x \rfloor$ the integer part of x. In a jammed configuration, a dimer must be followed either by another dimer or by a *single* vacant site. Thus a jammed configuration may be symbolically written as $\cdots$DDODDOD$\cdots$. That is, between each pair of dimers there may be either one vacancy or nothing. Each such string corresponds to a distinct jammed state. Since a vacancy can appear between any pair of dimers and also between a dimer and the end of the chain, there are $N+1$ possible locations for the $L-2N$ vacant sites. Therefore total number of distinct arrangements with N dimers is given by the binomial coefficient

$$F_{L,N} = \binom{N+1}{L-2N}, \tag{6.19}$$

and the total number of configurations with any number of dimers in the allowed range is $F_L = \sum_N F_{L,N}$.

In the thermodynamic limit, we fix the coverage $\rho = 2N/L$ and then evaluate $\ln F_{L,\rho}$ by keeping only the two leading terms in the Stirling formula $\ln x! \sim x \ln x - x$. In this approximation, the total number of fixed-density configurations also grows exponentially with the system size, $F_{L,\rho} \sim e^{Lf(\rho)}$, with

$$f(\rho) = \frac{\rho}{2}\ln\frac{\rho}{2} - (1-\rho)\ln(1-\rho) - \left(\frac{3\rho}{2}-1\right)\ln\left(\frac{3\rho}{2}-1\right). \tag{6.20}$$

Because $F_{L,\rho}$ grows exponentially with L, $f(\rho)$ is asymptotically dominated by its most probable value. Setting $f' = 0$ leads to $4\rho(1-\rho)^2 = (3\rho-2)^3$, whose solution gives the asymptotic equilibrium density $\rho_{\rm eq} = 0.823991\ldots$. Expanding f about $\rho_{\rm eq}$, the density dependence of the number of jammed configurations approaches the Gaussian

$$F_{L,N} \simeq \frac{F_L}{\sqrt{2\pi\Delta^2}}\, e^{-(\rho-\rho_{\rm eq})^2/2\Delta^2}), \tag{6.21}$$

with variance $\Delta = (Lf''(\rho_{\rm eq}))^{-1/2} \approx 0.261378 L^{-1/2}$, and prefactor fixed to give the correct total number of configurations.

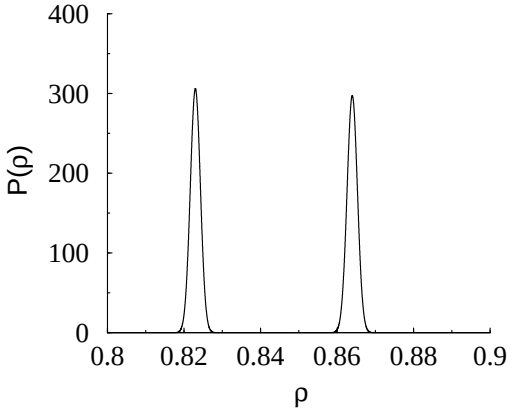


Figure 6.6: The equilibrium distribution (left) and jamming distribution (right) for a system of size $L = 10^4$.

The equilibrium probability distribution and the distribution of jammed states have the same functional form, except that their peak locations are different: $\rho_{\rm jam} \neq \rho_{\rm eq}$! If every jammed configuration had the same likelihood to occur, the jamming coverage should equal $\rho_{\rm eq} \approx 0.822991$ instead of $\rho_{\rm jam} \approx 0.864665$. Why are the jamming and the equilibrium distributions different? Equilibrium systems uniformly sample their phase space so that all microscopic configurations with the same energy are equiprobable. In contrast, for non-equilibrium systems, as manifested by irreversible adsorption, the dynamics dictates how the phase space is explored, and there is no reason that all microscopic configurations are sampled uniformly. Non-equilibrium systems need not minimize a free energy, nor explore all microscopic configurations equiprobably.

6.3 Correlations and Fluctuations

The relation $\rho = 1 - E_1$ is a simple example of expressing a fundamental physical quantity (the coverage) in terms of the empty interval probabilities E_m. As we now show, the empty interval probabilities contain much more information about the substrate occupancy, such as spatial correlation functions between occupied sites and fluctuations in surface coverage. Let's denote the probability of an arbitrary configuration by $\mathcal{P}[\cdots]$. Thus, for example, $E_2 = \mathcal{P}[\circ\circ]$, where again the states of the sites external to the string are not specified. With this notation, $\rho = \mathcal{P}[\bullet] = 1 - E_1$ is the consequence of the conservation statement $\mathcal{P}[\circ] + \mathcal{P}[\bullet] = 1$. Other conservation statements, such as

$$\mathcal{P}[\circ\circ] + \mathcal{P}[\circ\bullet] = \mathcal{P}[\circ] \quad \text{or} \quad \mathcal{P}[\circ\circ] + \mathcal{P}[\circ\bullet] + \mathcal{P}[\bullet\circ] + \mathcal{P}[\bullet\bullet] = 1$$

express the probability of any configuration of occupied sites in terms of probabilities of empty configurations. For simple configurations, these probabilities can be expressed only in terms of the empty interval probabilities E_m. For example, using the conservation statement

$$\mathcal{P}[\underbrace{\circ\cdots\circ}_{m}\bullet] + \mathcal{P}[\underbrace{\circ\cdots\circ}_{m}\circ] = \mathcal{P}[\underbrace{\circ\cdots\circ}_{m+1}],$$

then the probability for a configuration with at least m empty sites followed by a filled site is

$$\mathcal{P}[\underbrace{\circ\cdots\circ}_{m}\bullet] = E_m - E_{m+1}, \tag{6.22}$$

while the probability to find a void of length exactly equal to m is

$$V_m = \mathcal{P}[\bullet\underbrace{\circ\cdots\circ}_{m}\bullet] = E_m - 2E_{m+1} + E_{m+2}. \tag{6.23}$$

A fundamental characterization of correlations between occupied sites is the *pair correlation function* C_j

$$C_j \equiv \langle n_0\, n_j\rangle - \langle n_0\rangle\langle n_j\rangle. \tag{6.24}$$

Here n_j is the density at site j and the term $\langle n_0\, n_j\rangle$ may be graphically represented as

$$\langle n_0\, n_j\rangle = \mathcal{P}[\bullet\underbrace{\times\cdots\times}_{j-1}\bullet],$$

where $\times$ denotes a site whose state is unspecified. As we now show, to determine correlation functions such as C_j, we need configurations that include disconnected empty configurations. We thus denote

$$E_{i,j,k} = \mathcal{P}\big[\underbrace{\circ\cdots\circ}_{i}\ \underbrace{\times\cdots\times}_{j-1}\ \underbrace{\circ\cdots\circ}_{k}\big],$$

as the probability for the configuration that consists of two empty clusters of at least i and at least k sites that surround $j-1$ sites of unspecified states. Then $E_{i,1,k} = E_{i+k}$, where the latter quantity is the empty interval probability. For dimer adsorption, notice also that $E_{i,2,k} = \mathcal{P}[\circ\cdots\circ\times\circ\cdots\circ] = E_{i+1+k}$, since a *single* site that is sandwiched between clusters of empty sites must also be empty.

The probabilities $E_{i,j,k}$ satisfy a hierarchy of master equations similar to (6.2):

$$\begin{aligned}\frac{dE_{i,j,k}}{dt} =& -(i-1)E_{i,j,k} - E_{i+1,j,k} - E_{i+1,j-1,k}\\ &-(k-1)E_{i,j,k} - E_{i,j,k+1} - E_{i,j-1,k+1},\end{aligned} \tag{6.25}$$

for $i,k \geq 1$ and $j \geq 2$. The consecutive terms in the first line account for adsorption of a dimer within the empty i-string, overlapping the left end of this empty string, and overlapping the right end of this string; the terms in the second line are counterparts for the empty k-string. To solve this master equation, we generalize the ansatz (6.4) to:

$$E_{i,j,k}(t) = e^{-(i+k-2)t}\,\Psi_j(t) \quad \text{for} \quad i,k \geq 1, \tag{6.26}$$

where $\Psi_j(t) \equiv E_{1,j,1}$, to simplify the above master equations to:

$$\frac{d\Psi_j}{dt} = -2e^{-t}[\Psi_j + \Psi_{j-1}], \tag{6.27}$$

for $j \geq 2$, while for $j = 1$, $\Psi_1 = E_2$. Equations (6.27) are recursive and solvable by introducing the generating function $\Psi(x,t) = \sum_{j=2}^{\infty} x^j\, \Psi_j(t)$ to recast (6.27) into

$$\frac{\partial \Psi(x,t)}{\partial t} = -2e^{-t}\left[(1+x)\Psi(x,t) + x^2\, E_2(t)\right]. \tag{6.28}$$

The initial condition is $\Psi_j(t=0) = 1$, or $\Psi(x,0) = x^2(1-x)^{-1}$. Solving (6.28) subject to this initial condition and expanding the solution in a Taylor series in x we obtain (for $j \geq 2$):

$$\Psi_j = (E_1)^2 - E_1\left[\frac{(\ln E_1)^j}{2\cdot j!} + \sum_{k\geq j+1}\frac{(\ln E_1)^k}{k!}\right].$$

We now exploit these results to compute the pair correlation C_j in Eq. (6.24). Using the conservation statements

$$\begin{aligned}
&\mathcal{P}[\bullet\times\cdots\times\bullet] + \mathcal{P}[\bullet\times\cdots\times\circ] + \mathcal{P}[\circ\times\cdots\times\bullet] + \mathcal{P}[\circ\times\cdots\times\circ] = 1,\\
&\mathcal{P}[\circ\times\cdots\times\bullet] + \mathcal{P}[\circ\times\cdots\times\circ] = \mathcal{P}[\circ] = E_1,
\end{aligned}$$

we have $\langle n_0 n_j\rangle = 1 + \Psi_j - 2E_1$. Since $\langle n\rangle = 1 - E_1$, we finally obtain $C_j = \Psi_j - (E_1)^2$. Explicitly, the correlation functions are:

$$C_1 = E_2 - (E_1)^2, \qquad C_j = -E_1\left[\frac{(\ln E_1)^j}{2\cdot j!} + \sum_{k\geq j+1}\frac{(\ln E_1)^k}{k!}\right] \quad j > 1. \tag{6.29}$$

In the jammed state $E_1(\infty) = e^{-2}$, so that the limiting value of the pair correlation is

$$C_j \to -e^{-2}\frac{(-2)^{j-1}}{j!} \quad \text{as} \quad j\to\infty.$$

This super-exponential decay is much faster than typical exponential decay of correlations in many equilibrium systems with short-range interactions, such as a system of hard disks.

6.4 Adsorption in Higher Dimensions

Most applications of irreversible adsorption involve two-dimensional substrates. It is natural to begin with the irreversible adsorption of elementary objects such as disks, squares, rectangles, and sticks as a prelude to real systems, such as proteins and latex particles. To get a feeling for numbers, the jamming coverages for random sequential adsorption of various elementary objects in two dimensions are listed in Table 6.2. These coverages strongly depend on the shape of the object. An exact analysis of adsorption is generally not possible in higher dimensions, and one has to resort to approximations and numerical simulations. The one-dimensional theory still serves as a useful guide, however, because the evolution of the coverage has the same qualitative features in arbitrary dimensions. Thus, for example, the relaxation is exponential on discrete substrates and algebraic on continuous ones. Moreover, fluctuations in the number of adsorbed particles are extensive, *i.e.*, proportional to the volume. Finally, different jammed configurations are realized with different likelihoods, *i.e.*, the dynamics is non ergodic.

Discrete substrates

Adsorption *is* exactly soluble for one special high-dimensional substrate — the Bethe lattice, in which each site is connected to exactly z other sites in tree structure (Fig. 6.7). For dimer adsorption on the Bethe

object	substrate	$\rho_{\rm jam}$
unoriented dimers	square lattice	0.9068
2×2 squares	square lattice	0.74788
(aligned) squares	continuum	0.562009
disks	continuum	0.5472

Table 6.2: Jamming coverages for various objects in two dimensions.

lattice, the fundamental quantities are not empty intervals of length $\geq m$, but rather connected *clusters* of $\geq m$ sites that we again write as E_m. Because these clusters have a tree structure, it is easy to count the "boundary" configurations that enter into the master equations.

The probability E_m that all sites in such a cluster remain vacant during adsorption of dimers satisfies the master equation [compare with Eq. (6.2)]

$$\frac{dE_m}{dt} = -(m-1)E_m - [(z-2)m+2]E_{m+1}, \tag{6.30}$$

for $m \geq 1$, with the initial conditions $E_m(0) = 1$. The first term on the right accounts for deposition events in which the dimer lands somewhere within the cluster. The number of available locations for such "interior" adsorption events is just the number of bonds in the cluster. Since any cluster has a tree topology, the number of bonds is just $m-1$. The second term accounts for adsorption events in which the dimer lands with one site in the cluster and one site outside. The number of ways that such an event can occur equals the number of cluster perimeter sites — sites that adjoin the cluster, but are not part of the cluster itself (Fig. 6.7). For a cluster of 2 sites, the number of perimeter sites is $2(z-1)$. When a site is added to the cluster, 1 perimeter site is lost, but $(z-1)$ perimeter sites are gained. Continuing this counting for a cluster of m sites, the number of perimeter sites is $zm - 2(z-1)$.

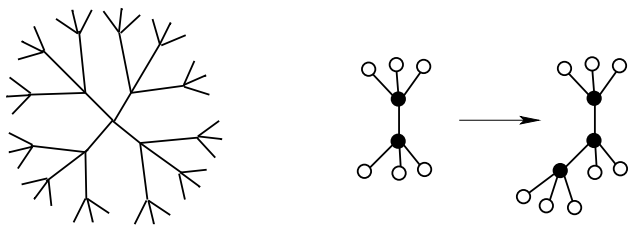


Figure 6.7: (Left) First three generations of a Bethe lattice with coordination number $z = 4$. (Right) Counting perimeter sites (circles) starting with a connected cluster of $m = 2$ and 3 sites (dots).

To solve the master equation (6.30), we again apply the exponential ansatz $E_m(t) = [\varphi(t)]^{m-1}\,\Phi(t)$ with $\varphi(0) = \Phi(0) = 1$ to match the initial condition (see the box on the next page). With this ansatz, the hierarchy of rate equations reduces to the two coupled differential equations

$$\frac{d\varphi}{dt} = -\varphi - (z-2)\varphi^2 \qquad \frac{d\Phi}{dt} = -z\,\varphi\,\Phi,$$

whose solutions are $\varphi = e^{-t}[(z-1)-(z-2)e^{-t}]^{-1}$ and $\Phi(t) = [(z-1)-(z-2)e^{-t}]^{-z/(z-2)}$. Consequently, the empty cluster probabilities are

$$E_m(t) = e^{-(m-1)t}\left[(z-1)-(z-2)\,e^{-t}\right]^{-m-2/(z-2)}. \tag{6.31}$$

The approach to the jamming coverage is exponential in time, with the jamming coverage equal to

$$\rho_{\rm jam} = 1 - (z-1)^{-z/(z-2)}. \tag{6.32}$$

In the limit $z \downarrow 2$, we recover the one-dimensional result, while for large z, the uncovered fraction is inversely proportional to the coordination number, $1-\rho_{\rm jam} \sim z^{-1}$. Amusingly, the Bethe lattice provides a good

approximation for $\rho_{\rm jam}$ for a regular lattice with the same coordination number. For example, when $z = 4$, dimer adsorption on the Bethe lattice gives $\rho_{\rm jam} = 8/9$, while for the square lattice, $\rho_{\rm jam} \approx 0.9068$.

A second look at almost exponential solutions

Consider a master equation with the generic form

$$\frac{dE_m}{dt} = \lambda(m+\alpha)E_m + \mu(m+\beta)E_{m+1}, \tag{6.33}$$

that encompasses Eq. (6.30). Again, the almost exponential ansatz provides an easy route to the solution. Let's assume a solution of the form

$$E_m = \Phi(t)[\varphi(t)]^m. \tag{6.34}$$

Substituting into Eq. (6.33) and then dividing by E_m gives

$$\frac{\dot\Phi}{\Phi} + m\frac{\dot\varphi}{\varphi} = \lambda(m+\alpha) + \mu(m+\beta)\varphi.$$

This result shows the utility of the ansatz (6.34), as the above equation divides naturally into terms linear in m and terms independent of m.

From the terms linear in m, we have

$$\frac{\dot\varphi}{\varphi} = \lambda + \mu\varphi, \tag{6.35}$$

from which we obtain $\varphi(t)$. The m-independent terms give

$$\frac{\dot\Phi}{\Phi} = \lambda\alpha + \mu\beta\varphi, \tag{6.36}$$

which then gives $\Phi(t)$, after which the original problem is solved.

The crucial point is that the factor φ^m in the original ansatz separates the initial set of equations (6.34) into two equations: one linear in m and one independent of m.

For discrete substrates in arbitrary spatial dimension, we can only give a heuristic argument that the relaxation to the jamming coverage decays exponentially in time

$$\rho_{\rm jam} - \rho(t) \sim e^{-\lambda t}\,. \tag{6.37}$$

As a concrete example, consider the dimer adsorption on the square lattice. At long times, the available spaces that can accommodate additional dimers are few and far between. These "target" regions are small clusters of unoccupied sites: dimers, trimers (both linear and bent), 2×2 squares, *etc.* To determine the rate at which these "lattice animals" get filled, we need the probabilities that these various configurations are empty. Crucially, the probability to find a vacant cluster on the substrate rapidly decreased with its size at long times. Thus only the smallest possible empty lattice animals persist and their asymptotic decay is dominated by the adsorption of dimers *inside* the animal. Thus for dimer adsorption on the square lattice, the probabilities of the simplest configurations (dimers, trimers, and 2×2 squares) evolve according to

$$\frac{d}{dt}\,P[\circ\circ] \sim -P[\circ\circ], \qquad \frac{d}{dt}\,P[\circ\,\circ\,\circ] \sim -2\,P[\circ\,\circ\,\circ], \qquad \frac{d}{dt}\,P\begin{bmatrix}\circ\circ\\ \circ\circ\end{bmatrix} \sim -4\,P\begin{bmatrix}\circ\circ\\ \circ\circ\end{bmatrix}.$$

Here, we use the shorthand $P[\cdot]$ to denote the likelihood of a configuration, and the numerical prefactor counts the number of ways that a dimer can adsorb within the cluster. The time dependence of these configurations therefore evolve as

$$P[\circ\circ] \sim e^{-t}\,, \qquad P[\circ\,\circ\,\circ] \sim e^{-2t}\,, \qquad P\begin{bmatrix}\circ\circ\\ \circ\circ\end{bmatrix} \sim e^{-4t}. \tag{6.38}$$

Generally, the probability that a given lattice animal is empty decays exponentially in time, $P(t) \sim \exp(-\lambda t)$, where λ counts the number of ways that a dimer can adsorb within a particular lattice animal. In particular, the coverage is determined by the rate equation $d\rho/dt \sim -2P[\circ\circ]$, so that

$$\rho_{\rm jam} - \rho(t) \sim e^{-t}. \tag{6.39}$$

A similar exponential relaxation arises for the adsorption of arbitrarily-shaped objects on discrete substrates in any dimension.

Continuous substrates

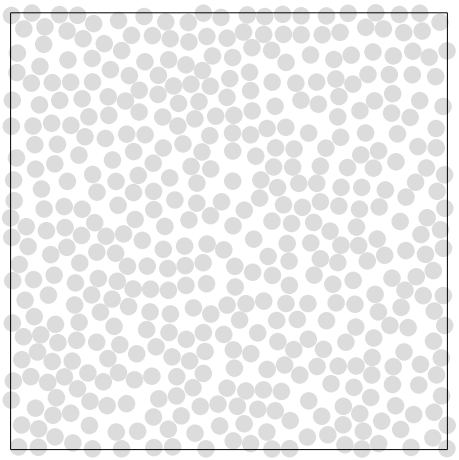


Figure 6.8: A jammed state for random sequential adsorption of disks in two dimension.

On continuous substrates, gaps between adjacent adsorbed objects can be arbitrarily small, and this feature leads to a slow algebraic relaxation of the density to the jamming density in which $\rho_{\rm jam} - \rho(t) \sim t^{-\sigma}$. For car parking in one dimension, we already demonstrated that $\sigma = 1$. Let's derive the corresponding decay for the adsorption of disks in two dimensions (Fig. 6.8). As the substrate approaches jamming, there will be only a small number of tiny and independent "target zones" within which the center of another disk can adsorb. To characterize these target zones, notice that around each disk there is an "exclusion zone" whose radius is twice that of the disk. An incident disk whose center lies within the exclusion zone of any already adsorbed disk cannot adsorb. The target zones of the substrate are the complement of the exclusion zones (Fig. 6.9). In a jammed configuration, no target zones remain even though the adsorbed particles do not completely cover the substrate.

Let $c(\ell, t)$ be the density of target zones of linear size ℓ. Because the area of such a target zone is quadratic in ℓ, the density of targets of linear size ℓ obeys $dc/dt \propto -\ell^2 c$, leading to the exponential decay

$$c(\ell, t) \sim e^{-\ell^2 t}. \tag{6.40}$$

Since each disk has the same area, the deviation of the substrate coverage from its jamming value is just proportional to the area fraction of the target zones:

$$\rho_{\rm jam} - \rho(t) \sim \int_0^\infty c(\ell, t)\, d\ell \sim \int_0^\infty e^{-\ell^2 t}\, d\ell \sim t^{-1/2}. \tag{6.41}$$

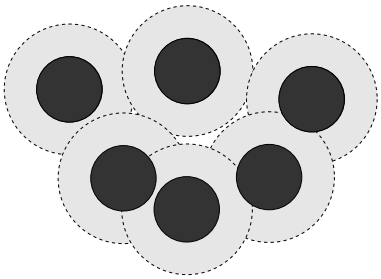


Figure 6.9: Two target areas (white), the exclusion zones (shaded), and the adsorbed disks (dark) near jamming. In the long-time limit only arc-triangular target areas, such as the tiny one on the left, remain.

Because target zones can be arbitrarily small, a power-law decay arises after rescaling the integral. Thus although the probability to find a target zone of a given size vanishes exponentially with time, the average over the sizes of all target zones leads to a power-law tail. This heuristic approach can be straightforwardly extended to arbitrary spatial dimension d. Now the area of a target zone of linear dimension ℓ scales as ℓ^d. Correspondingly, the density of target zones of linear dimension ℓ scales as $c(\ell, t) \sim e^{-\ell^d t}$. The analog of Eq. (6.41) then gives $\rho_{\rm jam} - \rho(t) \sim t^{-1/d}$ in d dimensions.

We can extend further this argument to determine the approach to jamming for elongated particles, for example, ellipses with a large aspect ratio. Now the notion of a target zone is no longer precise because minimum separation between two adsorbed ellipses depends on their relative orientations. However, as target zones get filled, there is a tendency for a newly-adsorbed ellipse to be oriented with its neighbors. This restriction plays an important role in the adsorption of non-symmetrical objects at long times. Let θ be the range of allowed orientations for an ellipse that is incident on a target zone of linear dimension ℓ. The density of target zones of linear size ℓ will asymptotically evolve according to $\frac{dc}{dt} \propto -\theta\, \ell^2\, c$. Since the orientational range vanishes in the long-time limit, we make the simplest self-consistent assumption that $\theta = a_1\ell + a_2\ell^2 + \cdots$. Then using $\theta \propto \ell$ for small ℓ, the concentration of target areas of linear dimension ℓ is governed by $dc/dt \propto -\ell^3\, c$. As a result,

$$c(\ell, t) \sim e^{-\ell^3 t}\,. \tag{6.42}$$

Substituting this form for $c(\ell, t)$ into Eq. (6.41), the relaxation to the jamming coverage is now given by $\rho_{\rm jam} - \rho(t) \sim t^{-\sigma}$, with $\sigma = 1/3$. Thus the orientational constraint hinders the approach to the jammed state.

For oriented squares, the target areas asymptotically are rectangular and the density $n(x, y, t)$ of target rectangles of size $x \times y$ decays according to $\frac{dc}{dt} = -xy\, c$. Consequently, $c \sim e^{-xy\, t}$, from which

$$\begin{aligned}\rho_{\rm jam} - \rho(t) = \int_0^1 \int_0^1 n(x, y, t)\, dx\, dy &\sim \int_0^1 \frac{1 - e^{-xt}}{xt}\, dx \\ &= t^{-1} \int_0^t \frac{1 - e^{-u}}{u}\, du \\ &\simeq t^{-1} \ln t. \end{aligned} \tag{6.43}$$

Similarly, the approach to jamming for the random sequential adsorption of aligned hypercubes in d dimensions is given by

$$\rho_{\rm jam} - \rho(t) \sim (\ln t)^{d-1}\, t^{-1}. \tag{6.44}$$

Needles

An particularly intriguing example is the deposition of zero-area, unoriented, identical needles of unit length — the limit of ellipses with a diverging aspect ratio. Here the areal coverage of the substrate vanishes, even though the number of adsorbed particles diverges with time. Early deposition attempts are mostly

successful because the substrate is nearly empty. Thus the number of adsorbed needles starts growing linearly with time and these initial needles have little orientational or positional order. However, when the needle density becomes of the order of 1, previously adsorbed needles strongly constrain both the position and the orientation of subsequent adsorption events. In the long time limit, domains form in which neighboring needles are nearly aligned *and* positionally ordered (Fig. 6.10). The orientation of each domain is fixed by the first few adsorption attempts, so that there is orientational order at small scales but not at large scales. Once domains are well defined, most adsorption attempts fail and the number of adsorbed needles grows sub-linearly with time.

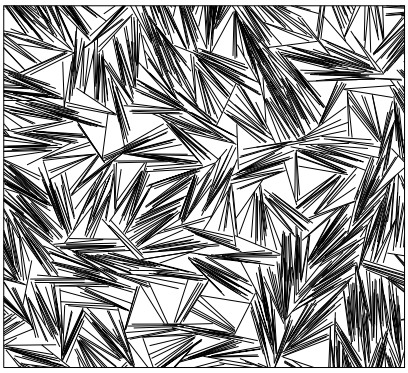


Figure 6.10: Random sequential adsorption of needles in two dimensions.

The formation of nearly aligned domains considerably simplifies the analysis of the late stages of adsorption. For a new needle to adsorb between two nearby needles, it must be closely aligned with them, both orientationally and positionally. The target area is approximately a unit-height trapezoid, with base widths x_1 and $x_2 \approx x_1$. A new needle may adsorb with its bottom end in the range $0 < y_1 < x_1$ and its top end in the range $0 < y_2 < x_2$ (Fig. 6.11). Such an adsorption event divides the trapezoid into two smaller trapezoids with base widths y_1, y_2 and $x_1 - y_1, x_2 - y_2$; this defines a geometric fragmentation process similar to that discussed in Sec. 5.4. We now apply the techniques of that presentation for the needle problem.

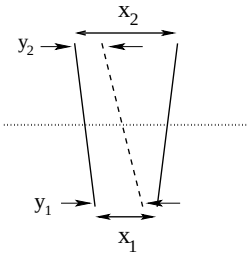


Figure 6.11: Random sequential adsorption of needles in one dimension as the fragmentation of trapezoids.

According to this representation of adsorption as trapezoid fragmentation, the density $c(x_1, x_2)$ of trapezoids with widths $\{x_1, x_2\}$ obeys the master equation

$$\frac{\partial}{\partial t} c(x_1, x_2) = -x_1 x_2 \, c(x_1, x_2) + 2 \int_{x_1}^{\infty} \int_{x_2}^{\infty} c(y_1, y_2) \, dy_1 \, dy_2. \tag{6.45}$$

The loss term is proportional to the total number of ways for the two ends of the needle to be placed along the bases of the trapezoid. The gain term accounts for the 2 ways in which the break-up of a larger trapezoid creates a trapezoid with base widths x_1, x_2.

The Mellin transform $M(s_1,s_2)=\int\int x_1^{s_1-1}x_2^{s_2-1}c(x_1,x_2)\,dx_1\,dx_2$ thus evolves according to

$$\frac{\partial}{\partial t}M(s_1,s_2)=\left(\frac{2}{s_1s_2}-1\right)M(s_1+1,s_2+1). \tag{6.46}$$

As in rectangular fragmentation, there is an infinite family of hidden conservation laws defined by $s_1^*s_2^*=2$. Assuming that the Mellin transform has the algebraic time dependence $M(s_1,s_2)\sim t^{-\alpha(s_1,s_2)}$, then Eq. (6.46) gives the recursion $\alpha(s_1+1,s_2+1)=\alpha(s_1,s_2)+1$. Using this recursion and the condition $\alpha(s_1^*,s_2^*)=0$ along the parabola $s_1^*s_2^*=2$, the exponent $\alpha(s_1,s_2)$ may be determined by the same reasoning as that given in Sec. 5.4, and the result is

$$\alpha(s_1,s_2)=\frac{s_1+s_2-\sqrt{(s_1-s_2)^2+8}}{2}\,. \tag{6.47}$$

One striking consequence of this formula is that the number density of needles $n(t)=M(1,1)$ varies sublinearly in time:

$$n(t)\sim t^{\nu}\qquad\text{with}\quad \nu=\sqrt{2}-1. \tag{6.48}$$

Another basic aspect of needle adsorption is their increasing degree of alignment. We quantify this alignment by $\langle \mathbf{n}_i\cdot\mathbf{n}_{i+1}\rangle=\langle\cos\theta\rangle$, with $\mathbf{n}_i$ a unit vector parallel to ith needle and θ the angle between the two adjacent needles. This angle is related to the base length of the trapezoid defined by two needles via $\theta=|x_1-x_2|$ for $x_1,x_2\ll 1$. Thus the orientational correlation function $\langle\cos\theta\rangle$ is related to the moments through $1-\langle\cos\theta\rangle\sim\langle(x_1-x_2)^2\rangle=2(\langle x^2\rangle-\langle x_1x_2\rangle)$, with $\langle x_1^2\rangle=\langle x_1^2\rangle\equiv\langle x^2\rangle$. Using $\langle x^2\rangle=M(3,1)/M(1,1)\sim t^{-(\alpha(3,1)-\alpha(1,1))}$, the orientational correlation function is then

$$1-\langle\cos\theta\rangle\sim t^{-\mu}\qquad\text{with}\qquad \mu=1+\sqrt{2}-\sqrt{3}=0.682162\ldots. \tag{6.49}$$

Thus neighboring needles asymptotically become perfectly aligned and the relaxation to this aligned state decays algebraically with time.

While the connection between needle adsorption and rectangular fragmentation involves some leaps of faith and hard-to-justify approximations, we gain in finding a natural way to account for the subtle multiscaling and non-rational exponents that describe needle adsorption.

6.5 Extensions

Thus far, we've focused on irreversible adsorption — once a particle adsorbs, it is immobile. Furthermore, we tacitly assumed that the only interaction is geometrical exclusion in which the adsorption probability depends only on whether there exists sufficient empty space to accommodate an incoming particle, and not on the distance to previously-adsorbed particles. Both of these assumptions are idealizations of reality, however, and we now study physically-motivated extensions of adsorption in which these assumptions are relaxed.

Cooperative Monomer Adsorption

A simple example of a distance-dependent interaction is the irreversible adsorption of monomers in one dimension in which adsorption is forbidden at sites next to already-occupied sites. Suppose that adsorption at an empty site occurs with rate 1 if both neighbors are empty, with rate r if exactly one neighbor is occupied, and rate 0 if both neighbors are occupied. The approach to jamming can again be treated in terms of empty interval probabilities $E_m(t)$. These probabilities now obey the master equations [compare with Eqs. (6.2)]

$$\begin{aligned}\frac{dE_1}{dt}&=-2rE_2-(1-2r)E_3 && m=1\\ \frac{dE_m}{dt}&=-(m-2+2r)E_m-2(1-r)E_{m+1} && m\geq 2.\end{aligned} \tag{6.50}$$

The first equation accounts for the disappearance of an empty interval of length 1, an event that occurs with probability $\mathcal{P}[\circ\,\circ\,\circ]+2r\mathcal{P}[\bullet\,\circ\,\circ]$. Here $\mathcal{P}[\circ\,\circ\,\circ]$ accounts for the loss of an empty site that is surrounded by

empty sites, while $2r\mathcal{P}[\bullet\circ\circ]$ accounts for the loss of an empty site in which one of its neighbors is already occupied. We then use $\mathcal{P}[\circ\circ\circ] = E_3$ and $\mathcal{P}[\bullet\circ\circ] = E_2 - E_3$ [see Eq. (6.22)] to give the first equation. The second equation accounts for the disappearance of an m-site empty interval. In such an interval, the particle can adsorb at $m-2$ sites in which both neighbors are empty (the factor $(m-2)E_m$). With probability $2r\mathcal{P}[\bullet\underbrace{\circ\circ\cdots\circ}_{m}]$ the monomer adsorbs next to an occupied site, and with probability $2\mathcal{P}[\underbrace{\circ\circ\cdots\circ\circ}_{m+1}]$, the monomer adsorbs one site in from the edge and an empty interval of length m is eliminated. Using $\mathcal{P}[\bullet\underbrace{\circ\circ\cdots\circ}_{m}] = E_m - E_{m+1}$[Eq. (6.22)], we obtain the second of Eqs. (6.50).

For $m \geq 2$, we again seek an exponential solution of the form $E_m(t) = e^{-(m-2+2r)t}\Phi(t)$, similar to Eq. (6.4). Substituting this ansatz into (6.50) gives

$$\frac{d\Phi}{dt} = -2(1-r)\,e^{-t}\,\Phi,$$

whose solution, subject to the initial condition $\Phi(0) = 1$, is

$$\Phi(t) = \exp\left[-2(1-r)\left(1-e^{-t}\right)\right].$$

Finally, by substituting $E_2(t) = e^{-2rt}\,\Phi(t)$ and $E_3(t) = e^{-(1+2r)t}\,\Phi(t)$ into the first of (6.50) and integrating we obtain

$$E_1(t) = 1 - \int_0^t \Phi(u)\left[2re^{-2ru} + (1-2r)\,e^{-(1+2r)u}\right]\,du.$$

Hence the surface coverage evolves as

$$\rho(r;t) = \int_0^t \left[2r + (1-2r)\,e^{-u}\right]\,e^{-2ru-2(1-r)(1-e^{-u})}\,du. \tag{6.51}$$

An amazing aspect of this solution is that the behavior for $r = 0$ is not the same as the behavior in the limit $r \to 0$! When $r = 0$, the above integral gives $\rho_{\rm jam}^0 \equiv \rho(0; t=\infty) = (1-e^{-2})/2$, which is one-half the jamming coverage in dimer adsorption. This result has an intuitive explanation that is illustrated graphically in Fig. 6.12. We define a dual lattice in which each site is halfway between the sites on the original lattice. Then each adsorbed monomer in the cooperative monomer problem is mapped onto an adsorbed dimer in the corresponding dimer system. Because of this one-to-one mapping between the two problems and also because each dimer occupies only a single site in the original lattice, the jammed density is simply one-half that of the dimer problem, namely, $(1-e^{-2})/2$.

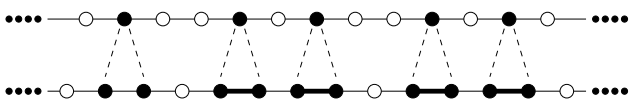


Figure 6.12: Top line: a jammed configuration for cooperative monomer adsorption when $r = 0$, where monomer adsorption requires that both neighbors are unoccupied. Bottom line: corresponding dimer configuration on the dual lattice.

On the other hand the final coverage for non-zero r, no matter how small is $\rho_{\rm jam}^+ \equiv \rho(r \to 0; \infty) = (1+e^{-2})/2$! Why is there a discontinuity in the jamming coverage for $r \to 0$? Physically, this discontinuity arises because of the wide separation of time scales in the two types of adsorption events that occur for infinitesimal r. In a time of the order of one, $r = 0$ adsorption events occur (adsorption at sites with both neighbors empty) until no further such events are possible. The system then reaches the $r = 0$ jammed state where empty sites can occur singly or in pairs. Then on a much longer time scale (of the order of r^{-1}), one site within each empty pair get filled, *e.g.*, $\bullet\circ\circ\bullet \to \bullet\bullet\circ\bullet$ or $\bullet\circ\circ\bullet \to \bullet\circ\bullet\bullet$. To determine $\rho_{\rm jam}^+$, consider first the jammed state for cooperative monomer adsorption. Let ρ_1 be the density of monomers that are followed by a single vacancy and ρ_2 the density of monomers that are followed by two consecutive

vacancies. By construction, $\rho_1 + \rho_2 = \rho_{\rm jam}^0 = (1 - e^{-2})/2$, and also $2\rho_1 + 3\rho_2 = 1$. Solving for ρ_2 gives $\rho_2 = e^{-2}$. After the final infill of one site in all consecutive vacancy pairs, the final jamming density is $\rho_{\rm jam}^+ = (1 - e^{-2})/2 + \rho_2 = (1 + e^{-2})/2$.

Mathematically, the source of the discontinuity is the first term in Eq. (6.51). In the limit $r \to 0$, the leading contribution of this term is

$$\int_0^\infty 2r\, e^{-2ru-2(1-e^{-u})}\, du.$$

When $r \to 0$, the main contribution to the integral is the region where u is large. In this regime, the factor e^{-u} in the exponent can be neglected and we are left with

$$\int_0^\infty 2r\, e^{-2ru-2}\, du = e^{-2}.$$

Thus for $r = 0^+$, the jamming density is $(1 - e^{-2})/2 + e^{-2} = (1 + e^{-2})/2$ as above.

Adsorbate Mobility

In irreversible adsorption, adsorbed molecules remain fixed where they first land on the substrate. In reality, particles find more stable positions, diffuse, desorb, *etc.* These post-adsorption events are often slow compared to the adsorption. We study the extreme limit where adsorption is infinitely faster than any post-adsorption event; for this limit, we may set the rate of post-adsorption processes to 1 and the adsorption rate to infinity. Thus whenever an adsorbate configuration permits another adsorption event, it occurs instantaneously. This separation of time scales for the two processes simplifies the analysis of this problem.

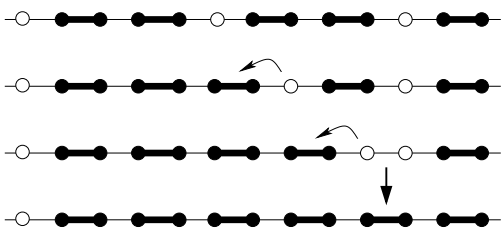


Figure 6.13: Example of mobile adsorbed dimers in one dimension. Each successive row shows the system after a single hopping event. When two holes become adjacent, they are immediately occupied by a dimer.

Suppose that dimers adsorb onto a one-dimensional lattice at an infinite rate and then undergo a simple random walk on this lattice if there is an empty space adjacent to the dimer to accommodate the dimer (overlapping of dimers is forbidden). Since the deposition rate is infinite, the lattice immediately reaches a quasi-jammed state in which all empty sites are isolated; we call such sites "holes". Once this state is reached, the dimers adjacent to holes can hop, as illustrated in Fig. 6.13. The hopping of a dimer to the left results in the effective hopping of the hole by 2 steps to the right.

Because the holes have a non-zero density, two adjacent holes may arise. When such a configuration occurs, it is instantaneously and irreversibly filled by a dimer. Consequently, the number of holes decreases and eventually every site of the lattice is occupied. This evolution is equivalent to the holes undergoing the diffusion-controlled binary annihilation,

$$H + H \to \emptyset,$$

whose dynamics is well-known (see chapter 9). Here the term diffusion-controlled signifies that diffusion controls the overall reaction rate, since annihilation occurs instantaneously whenever it is possible. Thus we infer that the density of holes decreases as

$$1 - \rho(t) \sim \begin{cases} t^{-1/2} & d = 1; \\ t^{-1}\, \ln t & d = 2; \\ t^{-1} & d > 2. \end{cases} \tag{6.52}$$

In writing the result for spatial dimension $d > 1$, we make the assumption that the hole motion is asymptotically diffusive when the lattice is nearly completely occupied by randomly -oriented dimers.

We can similarly analyze lattice deposition of trimers. Now holes hop by three lattice sites whenever a trimer hops by one site. When three holes are adjacent they undergo diffusion-controlled ternary annihilation,

$$H + H + H \to \emptyset,$$

whose dynamics is also known:

$$1 - \rho(t) \sim \begin{cases} t^{-1/2}\sqrt{\ln t} & d = 1; \\ t^{-1/2} & d > 1. \end{cases} \tag{6.53}$$

Here we have glossed over the issues of the shape of the trimer (straight or bent) and their orientation. It seems plausible that these details do not matter in the long-time limit when few holes remain. For the adsorption of diffusing M-mers, the long-time relaxation is controlled by the M-body annihilation

$$\underbrace{H + \cdots + H}_{M} \to \emptyset,$$

which leads to

$$1 - \rho(t) \sim t^{-1/(M-1)} \qquad M \geq 4. \tag{6.54}$$

Reversible car parking

In real adsorption processes, an adsorbed molecule has a finite binding energy to the substrate. Thus an adsorbed molecule will desorb at a rate that depends on the ratio of the binding energy to the temperature. If a fixed density of molecules is maintained in a gas phase above the substrate, we then have *reversible* adsorption-desorption: molecules adsorb with a rate k_+ and desorb with a rate k_-. While adsorption is subject to the availability of space, desorption events occur independently for each adsorbed molecule. It is fun to think of the monomers as cars that are trying to park along a one-dimensional curb; we ignore the fact that a car needs a little extra room to actually steer into a parking space (Fig. 6.14). From everyday experience, we all know it is hard to find a good parking spot. If the rate at which cars leave — the desorption rate — is slow, the probability to find a parking spot large enough to accommodate your car becomes very small. When the position of individual parking spots is unregulated, such as in "resident-only" parking areas, cars will typically be very tightly packed, and the car density approaches the limit of perfect packing as the desorption rate goes to zero. This feature is a surprising outcome of the competition between adsorption and desorption. While the steady-state coverage is nearly complete for infinitesimal desorption, the coverage is significantly less than complete for no desorption.

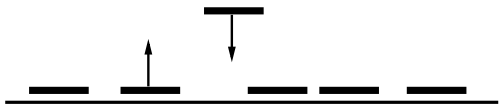


Figure 6.14: Adsorption-desorption of cars. Desorption attempts are always allowed, while the adsorption attempt shown fails because of insufficient parking space.

Another motivation for studying reversible adsorption is its connection to granular compaction. Suppose that identical glass beads are placed into a jar one at a time but randomly. The density of filled space in this bead packing — approximately 0.58 — is known as the *random packing density*. If this bead-filled jar is then vibrated vertically at a suitable intensity and frequency, the density will slowly increase and eventually relax to the *random close-packing density* of approximately 0.68. In experiments, the time scale over which this compaction occurs can be as long as months! Moreover, this random close packing-density is still smaller than the maximal packing fraction of $\pi/\sqrt{18} \approx 0.7405$ for a face-centered cubic lattice sphere pack. This compaction is analogous to adsorption-desorption. After the jar is initially filled, there exist many interstitial empty spaces that are not large enough to accommodate a bead. Because of the vibrations,

occasional local re-arrangements occur that gradually eliminate these empty spaces. As the density increases, re-arrangements become progressively more collective in nature and thus more rare. This slow compaction seems to be captured mathematically in terms of adsorption-desorption.

As a preliminary for the car parking problem, consider the trivial example of adsorbing and desorbing monomers that interact only with single sites on the substrate. The density of adsorbed particles ρ satisfies the Langmuir equation

$$\frac{d\rho}{dt} = -k_-\rho + k_+(1-\rho), \tag{6.55}$$

in which the total adsorption rate is proportional to the density of empty space and the total desorption rate is proportional to the density of adsorbed monomers. The time dependence of the density is

$$\rho(t) = \rho_\infty + (\rho_0 - \rho_\infty)\, e^{-t/\tau}, \tag{6.56}$$

with the relaxation time τ given by $\tau^{-1} = k_+ + k_-$ and the final coverage $\rho_\infty = k_+/(k_+ + k_-)$ Notice that as $h \equiv k_+/k_- \to \infty$, the asymptotic coverage is $\rho_\infty \approx 1 - h^{-1}$ is reached in a time $\tau \propto h^{-1}$.

For reversible car parking, it is more convenient to work with the density of voids of length x at time t, $V(x,t)$, rather than the density of empty intervals $E(x.t)$. As usual, the strategy is to write and then solve the governing master equations which are:

$$\frac{\partial V(x,t)}{\partial t} = \begin{cases} 2k_+ \displaystyle\int_{x+1}^{\infty} V(y,t)\, dy - 2k_- V(x,t) & x < 1; \\[2ex] 2k_+ \displaystyle\int_{x+1}^{\infty} V(y,t)\, dy - 2k_- V(x,t) - k_+(x-1)V(x,t) & \\ \qquad + \dfrac{k_-}{\int_0^\infty V(x,t)\, dx} \displaystyle\int_0^{x-1} V(y,t)V(x-y-1,t)\, dy & x > 1. \end{cases} \tag{6.57}$$

Each term has a simple explanation. For both $x < 1$ and $x > 1$, the first term on the right accounts for the creation of a void of length x when a car parks in a void of length $x+1$ or greater; the factor 2 accounts for the two places that the car can park in an $x+1$-void to create an x-void. The second term accounts for the loss of an x-void because of the desorption of a car at either end of the void. For $x > 1$, the third term accounts for the loss of an x-void when a car parks inside it. The last term for $x > 1$ is more subtle; it accounts for the creation of an x-void when a car leaves a parking spot that has an empty space of length y at one end of the car and a space $x-y-1$ at the other end. Thus a void of length x is created by merging voids of length y and $x-y-1$, together with the unit-length vacated parking spot. The correct way to express this composite event is through a 3-body correlation function. However, this description is not closed, as the void density is coupled to a 3-body function. We would then have to write an evolution equation for the 3-body correlation in terms of higher-body correlations, *ad infinitum*. To break this hierarchy at the simplest level, we invoke the mean-field assumption that the 3-body correlation function for a car to be flanked by voids of length y and $x-y-1$ is the product of single-void densities. The factor $\int V(x,t)\, dx$ in the denominator properly normalizes the probability that the neighbor of a y-void has length $x-y-1$.

Since there is a one-to-one correspondence between voids and adsorbed particles, the density of voids of any size equals the particle density; thus $\rho = \int V(x,t)\, dx$. Also, the mean fraction space occupied by voids and by cars equals 1; thus $1 = \int (x+1)\, V(x,t)\, dx$. Because each adsorption and desorption event changes the overall density by the same amount, the rate equation for the total density is simply

$$\frac{\partial \rho}{\partial t} = -k_-\rho + k_+ \int_1^\infty (x-1)\, V(x,t)\, dx, \tag{6.58}$$

which generalizes the Langmuir equation (6.55). The interpretation of this equation is straightforward: with rate $k_-\rho$, a parked car desorbs, while the second term accounts for the parking of a car in a space of length $x > 1$; this same equation can also be obtained by integrating the master equations (6.57) over all lengths.

Most of the interesting behavior about car parking can be easily gleaned by solving the master equations (6.57) in the steady state. Then the equation for $x < 1$ relates the void density to its spatial integral; this

fact suggests an exponential solution $V(x) = Ae^{-\alpha x}$. Substituting this ansatz into the the master equation for $x < 1$ gives the condition

$$h \equiv \frac{k_+}{k_-} = \alpha\, e^{\alpha}. \tag{6.59}$$

Next, applying the normalization condition $\int (x+1)\, V(x)\, dx = 1$ gives $A = \alpha^2/(\alpha+1)$. Thus the parked car density is simply

$$\rho = \int V(x)\, dx = \frac{\alpha}{\alpha+1}, \tag{6.60}$$

and eliminating α in favor of ρ, the probability distribution for parking spaces of length x is

$$V(x) = \frac{\rho^2}{1-\rho}\; e^{-\rho x/(1-\rho)}. \tag{6.61}$$

From Eqs. (6.59) and (6.60), the limiting behaviors of the steady-state density as a function of the scaled adsorption rate are:

$$\rho(h) \approx \begin{cases} h & h \to 0; \\ 1 - [\ln h]^{-1} & h \to \infty. \end{cases} \tag{6.62}$$

For slow adsorption, excluded-volume effects are negligible and the equilibrium density simply equals the adsorption rate, $\rho \approx h$. However, when the adsorption rate is large, excluded-volume effects are dominant. For example, to attain a steady-state density of $\rho = 0.99$, an astronomical ratio of adsorption to desorption rates, $h \approx e^{100}$, is needed! In contrast, for reversible monomer adsorption, a value of $h = 100$ gives a density of 0.99. As mentioned at the beginning of this section, a particularly intriguing feature is that the limiting behavior $\rho(h \to \infty) \to 1$ is distinct from the jamming density $\rho_{\rm jam} = \rho(h = \infty) = 0.747597\ldots$. The crucial point is that any infinitesimal desorption ($h \to \infty$ but still finite) eventually allows all wasted space to get filled. However, if $h = \infty$ at the outset, there is no mechanism to utilize too-small parking spaces.

Let's now study how the steady-state parking density is reached in the limit $h \to \infty$. For this purpose, the *quasi-static* approximation is extremely useful. In later chapters, we will see that this quasi-static approximation provides an easy route to solving a wide variety of slowly-varying phenomena. The basis of this approximation is the observation that car parking rarely succeeds as $h \to \infty$. Consequently, there is sufficient time for the voids to reach a nearly equilibrium state. Thus we use the steady-state exponential void density given by (6.61) in the rate equation (6.58). With this assumption, the density evolves as

$$\frac{d\rho}{dt} = -k_-\rho + k_+(1-\rho)\, e^{-\rho/(1-\rho)}. \tag{6.63}$$

The linear desorption term has the same form as in the monomer adsorption-desorption problem. However, the adsorption term is modified by the probability that an adsorption event is successful. This effective sticking probability $S(\rho) \equiv e^{-\rho/(1-\rho)}$ is extremely small when $\rho \to 1$, a feature that reflects the difficulty of finding a parking space when the density of parked cars is close to one.

We can obtain the effective sticking probability by the following heuristic argument. Consider a one-dimensional nearly-full parking lot with density $\rho = 1/(1+\langle x\rangle)$. Here $\langle x\rangle \ll 1$ represents the small average bumper-to-bumper distance between neighboring parked cars. Thus a driver who tries to park his car by happening upon a sufficiently large parking spot is almost always out of luck. Instead a driver has to enlist the help of $N = \langle x\rangle^{-1} = \frac{\rho}{1-\rho}$ owners of consecutive parked cars to move each of their cars forward a little; the first by $\langle x\rangle$, the second by $2\langle x\rangle$, the third by $3\langle x\rangle$, *etc.*, until a space of length one is generated. The probability of this cooperative rearrangement of picking cars sequentially and moving them forward by a prescribed amount decays exponentially in the number of cars. Thus the effective sticking probability $S \sim e^{-N} \sim e^{-\rho/(1-\rho)}$.

Finally, we determine the relaxation toward the steady state for $h \to \infty$. In this limiting case, the process is equivalent to a large population of cars are cruising the streets in a predatory manner, looking for parking spots. As soon as a parking spot becomes available it is immediately taken by a nearby cruising car. Technically, this situation is the *desorption-controlled* limit, as desorption limits of the overall reaction. Since any desorption event is immediately followed by at least one adsorption event, the effect of the loss term in Eq. (6.63) is effectively canceled out. Thus we obtain the time dependence of the density by neglecting the

loss term in (6.63). Because the gain term in this equation has been constructed from the steady-state void density, it remains positive for any $\rho < 1$ and therefore builds in a relaxation to the completely-occupied state. To solve the resulting rate equation $d\rho/d(k_+t) = (1-\rho)\exp[-\rho/(1-\rho)]$, we write $g = 1/(1-\rho)$ and approximate the equation by $dg/d(k_+t) \propto e^{-g}$ to yield the asymptotic behavior $g \sim \ln(k_+t)$. This gives an extremely slow logarithmic decay law

$$\rho(t) \sim 1 - (\ln k_+ t)^{-1}. \tag{6.64}$$

Lateral Adsorption

Our final example is adsorption with lateral relaxation, a process that can be recast as the computer science problem of a hash table construction. To illustrate the basis of the model, consider a gymnasium goer who daily uses a locker room in which all locker use is transient. Which locker should the user choose so that he won't forget his locker after his workout? The hash table approach is to start with a fixed number, say locker #123. If empty, take it. If occupied, move to #124 and take it, if it is free. If not, move to #125, *etc.*, until a vacancy is encountered. As long as the locker room is not too full, this approach provides a quick algorithm to find and remember a vacant locker reliably.

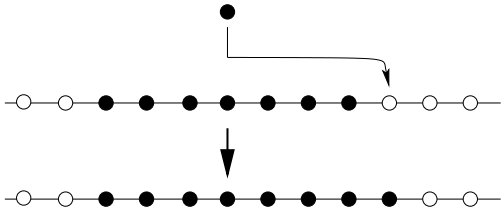


Figure 6.15: A monomer that is incident on an occupied site moves to the right and adsorbs when it first encounters a vacancy.

This search process can be viewed as the following adsorption problem. If an incoming particle is incident on an occupied part of the substrate, the particle moves laterally one site at a time along the substrate and absorbs when a vacancy is first encountered (Fig. 6.15). For convenience, we set the deposition rate equal to 1, while the subsequent search for a vacancy occurs at infinite rate. By construction, each deposition attempt is successful, so that the coverage $\rho(t) = t$ and the system is completely filled at $t = 1$. We now determine the distribution of voids as a function of time by studying the evolution of the empty interval probabilities $E_m(t)$.

The exact master equation for E_m involves

satisfy the approximate master equation

$$\frac{dE_m}{dt} = -mE_m - (E_m - E_{m+1})\,\langle m\rangle, \tag{6.65}$$

where $\langle m\rangle$ is the average size of occupied islands. The first term in this equation accounts for the direct deposition of a monomer into an empty m-site sequence. The second term accounts for adsorption in an m-interval in which the left boundary site occupied. Here the empty interval gets filled when a monomer is transported along a string of occupied sites until the empty interval is encountered. We have made the mean-field approximation that there is no correlation between the sizes of the island of occupied sites and the adjacent empty interval. With this assumption, the joint probability for an empty interval of length m and the size of the adjacent island may be written as a product of one-body quantities.

The master equation (6.65) still looks formidable because it does not appear to be closed—the mean value $\langle m\rangle$ is not expressed in terms of the E_m. However, there is a simple relation between $\langle m\rangle$ and E_m. By definition, the average island size is given by

$$\langle m\rangle = \frac{\sum mI_m}{\sum I_m}.$$

The sum in the numerator is just the density occupied sites, which simply equals t. The sum in the denominator is total island density I. Now the island density is the same as $\mathcal{P}(\circ\bullet)$, and the latter quantity is simple $E_1 - E_2$.

To solve the resulting master equation, we again attempt the exponential ansatz $E_m = e^{-mt}\,\Phi(t)$. Substituting this into (6.65), and using the above connection between $\langle m\rangle$ and $E_1 - E_2$ and t, we reduce the infinite set of differential equations (6.65) into $\dot\Phi = -t\,e^t$. It is striking that the empty interval method leads to a fortuitous set of cancellations that makes the problem amenable to solution. Integrating the equation for Φ, we obtain $\Phi = (1-t)\,e^t$, leading to the empty interval probabilities

$$E_m(t) = (1-t)\,e^{-(m-1)t}.$$

From this expression, the density of islands $I = E_1 - E_2$ and the density of voids $\mathcal{P}[\bullet\circ\cdots\circ\bullet] = E_m - 2E_{m+1} + E_{m+2}$ are given by:

$$\begin{aligned} I(t) &= (1-t)\,(1-e^{-t}) \\ V_m(t) &= (1-t)\,(1-e^{-t})^2\,e^{-(m-1)t}\,. \end{aligned}$$

A remarkable property of this problem is that in addition to obtaining the empty interval and empty void probabilities, we can also determine their filled analogs F_m and I_m, namely the filled interval and the island probabilities, respectively. These are defined by

$$F_m = \mathcal{P}[\underbrace{\bullet\cdots\bullet}_{m}], \qquad I_m = \mathcal{P}[\circ\underbrace{\bullet\cdots\bullet}_{m}\circ].$$

Let's compute the island densities. These densities evolve according to the master equations

$$\begin{aligned} \frac{dI_1}{dt} &= -\left(3 + \frac{tV_1}{I^2}\right) I_1 + \sum_{n=2}^{\infty}(n-2)V_n \\ \frac{dI_m}{dt} &= -\left(m+2+\frac{tV_1}{I^2}\right) I_m + \left(1-\frac{V_1}{I}\right)(m+1)I_{m-1} + \frac{V_1}{I^2}\sum_{n=1}^{m-2}(n+1)I_nI_{m-n-1} \qquad m\geq 2. \end{aligned} \tag{6.66}$$

(using the sum rules $\sum nV_n = 1-t$ and $\sum V_n = V$, the last term simplifies to $1-t-2V+V_1$). All terms in above equations are self-explanatory; the linear in densities terms are obviously exact while the nonlinear terms tacitly assume the lack of correlations between sizes of adjacent islands and voids. This subtle feature is also required in writing the second term on the right-hand side of Eq. (6.65).

Equations (6.66) are recursive and can be solved accordingly. For instance, we get $I_1 = (1-t)t\,e^{-2t}$ and $I_2 = \frac{3}{2}(1-t)t^2e^{-3t}$ leading to the conjectural behavior $I_m = a_m(1-t)t^me^{-(m+1)t}$. This ansatz indeed solves the problem if for $m\geq 2$ the amplitudes satisfy

$$ma_m = (m+1)a_{m-1} + \sum_{n=1}^{m-2}(n+1)a_na_{m-n-1}\,. \tag{6.67}$$

It is convenient to set $a_0 = 1$; then (6.67) holds for $m = 1$ since it gives $a_1 = a_0 = 1$. With the help of the generating $a(x) = \sum_{m\geq 0} a_m x^m$, we recast the recurrence (6.67) into the differential equation

$$\frac{da}{dx} = a^2 + xa\,\frac{da}{dx}\,. \tag{6.68}$$

Note that the scale transformation $x\to\lambda x$, $a\to\lambda^{-1}a$ leaves Eq. (6.68) unchanged thereby suggesting to use the scale invariant variable $y(x) = xa(x)$. The resulting equation $(1-y)y' = x^{-1}y$ is separable, and it is immediately solved to yield $x = y\,e^{-y}$. We sure can expand x in y but we must do the opposite and find

$y = \sum a_m x^{m+1}$. This is accomplished as follows:

$$\begin{aligned} a_m &= \frac{1}{2\pi i} \oint dx \, \frac{y(x)}{x^{m+2}} \\ &= \frac{1}{2\pi i} \oint dy \, \frac{y\, x'(y)}{[x(y)]^{m+2}} \\ &= \frac{1}{2\pi i} \oint dy \, \frac{(1-y)e^{(m+1)y}}{y^{m+1}} \\ &= \frac{(m+1)^{m-1}}{m!} \, . \end{aligned}$$

This completes the derivation of the density of islands:

$$I_m(t) = \frac{(m+1)^{m-1}}{m!} \, (1-t) \, t^m \, e^{-(m+1)t} \, .$$

The density of the empty strings is then found by integrating Eq. (6.69) twice subject to the given boundary values F_1 and F_2. We get

$$F_m = (m-1)F_2 - (m-1)F_1 + \sum_{k=1}^{m-2} (m-1-k) I_k,$$

with the first two values given by

$$\begin{aligned} F_1 &= 1 - E_1 = t, \\ F_2 &= 1 - 2E_1 + E_2 = 1 - 2(1-t) + (1-t)e^{-t} \, . \end{aligned}$$

6.6 Notes

The first equation (6.1) in this chapter was actually the first exact result in the field; it was derived in a pioneering work by Paul Flory (1939) in the context of the cyclization reaction on the polymer chain (his approach is essentially described in section 48). The jamming coverage for the car parking model was found by Alfred Rényi (1958) 49. In 60s, several people (particularly Cohen, Keller, Reiss, Widom) recognized the advantages of the evolution approach. Earlier exact results are reviewed and systematized in 50. More recent reviews 51; 52; 7; 53 are written by researches who were (and some still are) very active in the field in 80s and 90s. These reviews contain many useful things not covered in this chapter (e.g. multilayer adsorption — we discussed only monolayer case) and huge bibliography. Experimental techniques are reviewed by Ramsden 54.

The subject of sections 6.1–6.1 is classical although some of the results are quite recent and could not been found in reviews (e.g., models in which particles landing on the top of islands quickly diffuse to vacant sites 55). In sections 6.4 and 6.5 we gave a glimpse of adsorption problems for which there is currently no framework that allows to do analytical computations yet we already understand some interesting features, e.g. asymptotics. For adsorption on higher-dimensional substrates, the reader can like original papers that are short and lucid; see e.g. Refs. 56; 39 for discussion of (6.44). The deposition of needles is due to Tarjus and Viot 58. The connection with diffusion-controlled annihilation was recognized by Privman and co-workers (see 59), the parking lot model 60 is actively investigated due to success in explaining several features of granular materials.

Problems

Section 6.1

1. Compute the total density of voids and the density of islands.

2. Suppose that dimers adsorb onto a one-dimensional partly filled lattice. Consider the initial state where each lattice is independently occupied with probability ρ_0. Compute the jamming coverage.

3. Solve the car parking model if initially there is a density λ of randomly distributed sizeless defects.

Section 6.2

4. Suppose that dimers adsorb onto a one-dimensional lattice of length L. Starting with expression (6.16) for the generating function of the coverage, compute the average jamming coverage.

Section 6.3

5. Compute the magnitude of fluctuations in the number of adsorbed dimers in a region of L sites.

6. Compute the structure factor $S(q) \equiv \sum e^{iqm} C_m$.

7. Define the probabilities for filled strings of length m, F_m, and for islands of length m, I_m:

$$F_m \equiv \mathcal{P}[\underbrace{\bullet \cdots \bullet}_{m}], \qquad I_m \equiv \mathcal{P}[\circ \underbrace{\bullet \cdots \bullet}_{m} \circ].$$

Show that the island probability is the discrete second derivative of the filled string probability,

$$I_m = F_m - 2F_{m+1} + F_{m+2}, \tag{6.69}$$

8. For $m \leq 3$, use conservation statements to find the following relation between F_m and the empty interval probabilities E_m:

$$\begin{aligned} F_1 &= 1 - E_1, \\ F_2 &= 1 - 2E_1 + E_2, \\ F_3 &= 1 - 3E_1 + 2E_2. \end{aligned}$$

Notice that the general form of the last identity is

$$F_3 = 1 - 3\mathcal{P}[\circ] + 2\mathcal{P}[\circ\circ] + \mathcal{P}[\circ \times \circ] - \mathcal{P}[\circ \circ \circ],$$

but, for adsorption of dimers, $\mathcal{P}[\circ \times \circ] = \mathcal{P}[\circ \circ \circ]$ so that the 3-body terms in this identity cancel.

For $3 < m < 7$, express F_m in terms of E_j and the probability for two disconnected empty intervals $E_{i,j,k}$. For F_5, for example,

$$F_5 = 1 - 5E_1 + 4E_2 - 2E_4 + E_5 + 2(1 - e^{-t})\Psi_3 + \Psi_4.$$

Chapter 7

SPIN DYNAMICS

Kinetic spin systems play a crucial role in the development of non-equilibrium statistical physics. The prototypical example is the appealing-simple kinetic Ising model, in which the conventional Ising model of equilibrium statistical mechanics is endowed with physically-motivated transition rates that allows the system to "hop" between different microstates. Just as investigations of the equilibrium Ising model have elucidated the rich phenomenology underlying the transition between the disordered and ferromagnetically-ordered states, studies of the kinetic Ising model have yielded deep insights that have played a starring role in the development of the modern theory of critical phenomena and phase ordering kinetics.

7.1 The Voter Model

There is an even simpler kinetic spin system — the voter model — that will be the starting point for our discussion. One reason for focusing on the voter model first is that it is exactly soluble in all spatial dimensions. This solution also provides an instructive introduction for understanding kinetic Ising models. The voter model was introduced in the context of interacting particle systems and has been one of the most extensively examples of such systems. The voter model describes, in an appealing and paradigmatic way, how consensus emerges in a population of spineless individuals. That is, each individual has no firmly fixed opinion and merely takes the opinion of one of its neighbors in an update event. A finite population of such voters eventually achieves consensus in a time that depends on the system size and on the spatial dimension. In this section, we employ techniques inspired from non-equilibrium statistical physics, to solve some of the most basic and striking dynamical properties of the voter model on regular lattices in all dimensions.

In the voter model, an individuals is situated at each site of a graph. This graph could be a regular lattice in d dimensions, or it could be any type of graph—such as the Erdös-Rényi random graph, or a graph with a broad distribution of degrees. Each voter can be in one of two states that, for this presentation, we label as "Democrat" and "Republican". Mathematically, the state of the voter at $\mathbf{x}$, $s(\mathbf{x})$, can take the values ± 1 only; $s(\mathbf{x}) = +1$ for a Democrat and $s(\mathbf{x}) = -1$ for a Republican.

The dynamics of the voter model is simplicity itself. Each voter has no confidence and looks to a neighbor to decide what to do. A single update event in the voter model therefore consists of:

1. Pick a random voter.

2. The selected voter at $\mathbf{x}$ adopts the state of a randomly-selected neighbor at $\mathbf{y}$; that is, $s(\mathbf{x}) \to s(\mathbf{y})$.

3. Repeat steps 1 & 2 *ad infinitum* or stop when consensus is achieved.

Notice that a voter changes opinion only when its neighbor has the opposite opinion. A typical realization of the voter model on the square lattice is shown in Fig. 7.1, showing how the system tends to organize into single-opinion domains as time increases.

It is expedient to have each update step occur at a fixed rate. The rate at which a voter at $\mathbf{x}$ changes to

the state $-s(\mathbf{x})$ may then be written as

$$w(s(\mathbf{x})) = \frac{1}{2}\left(1 - \frac{s(\mathbf{x})}{z}\sum_{\mathbf{y}\ \text{n.n.}\mathbf{x}} s(\mathbf{y})\right), \tag{7.1}$$

where the sum is over the nearest neighbors of site $\mathbf{x}$. Here z is the coordination number of the graph and we tacitly assume that each site has the same coordination number. The basic feature of this dynamical rule is that the transition rate of a voter at $\mathbf{x}$ equals the fraction of disagreeing neighbors — when a voter at $\mathbf{x}$ and all its neighbors agree, the transition rate is zero; conversely, the transition rate equals 1 if all neighbors disagree with the voter at $\mathbf{x}$. This linearity is the primary reason why the voter model is soluble. One can generalize the voter model to include opinion changes, $s(\mathbf{x}) \to -s(\mathbf{x})$, whose rate does not depend on the local environment, by simply adding a constant to the flip rate.

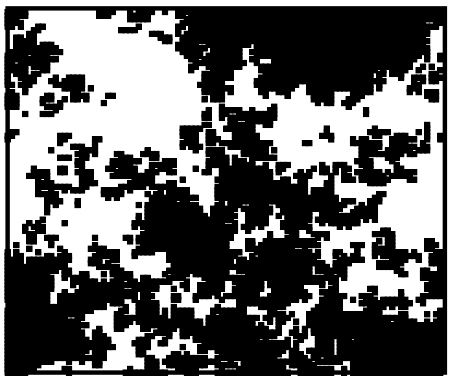


Figure 7.1: The voter model in two dimensions. Shown is a snapshot of a system on a 100×100 square lattice at time $t = 1000$, obtained by a Monte Carlo simulation. Black and white pixels denote the different opinion states.

To solve the voter model, we need, in principle, the probability distribution $P(\{\mathbf{s}\}, t)$ that the set of all voters are in configuration $\{\mathbf{s}\}$ at time t. This probability distribution satisfies the master equation

$$\frac{dP(\{\mathbf{s}\})}{dt} = -\sum_{\mathbf{x}} w(s(\mathbf{x}))P(\{\mathbf{s}\}) + \sum_{\mathbf{x}} w(-s(\mathbf{x}))P(\{\mathbf{s}\}_{\mathbf{x}}). \tag{7.2}$$

Here $\{\mathbf{s}\}_{\mathbf{x}}$ denotes the state that is the same as $\{\mathbf{s}\}$ except that the voter at $\mathbf{x}$ has changed opinion. In this master equation, the loss term accounts for all possible transitions out of state $\{\mathbf{s}\}$, while the gain term accounts for transitions to the state $\{\mathbf{s}\}$ from states in which one spin differs from the configuration $\{\mathbf{s}\}$. In principle, we can use this master equation to derive closed equations for all moments of the probability distribution — namely, all multi-spin correlation functions of the form $S_{\mathbf{x},\ldots,\mathbf{y}} \equiv \langle s(\mathbf{x})\cdots s(\mathbf{y})\rangle$ where the angle brackets denote the average $\langle f(\{\mathbf{s}\})\rangle \equiv \sum_{\mathbf{s}} f(\{\mathbf{s}\})P(\mathbf{s})$.

Let's begin by considering the simplest such correlation function, namely, the mean spin, or equivalently, the one-point function, $S(\mathbf{x}) \equiv \langle s(\mathbf{x})\rangle$. While it is possible to obtain the evolution of the mean spin and indeed any spin correlation function directly from the master equation (7.2), this approach involves some bookkeeping that is prone to error. We therefore present a simpler alternative method. In a small time interval Δt, the state of a given voter changes as follows:

$$s(\mathbf{x}, t+\Delta t) = \begin{cases} s(\mathbf{x},t) & \text{with probability } 1 - w(s(\mathbf{x}))\Delta t, \\ -s(\mathbf{x},t) & \text{with probability } w(s(\mathbf{x}))\Delta t. \end{cases} \tag{7.3}$$

Since the opinion at $\mathbf{x}$ changes by $-2s(\mathbf{x})$ with rate $w(s(\mathbf{x}))$, the average opinion evolves according to the rate equation

$$\frac{dS(\mathbf{x})}{dt} = -2\langle s(\mathbf{x})w(s(\mathbf{x}))\rangle. \tag{7.4}$$

Substituting in the transition rate from (7.1) and using the fact that $s(\mathbf{x})^2 = 1$, we find that for voters that are located on the sites of a d-dimensional hypercubic lattice, the rate equation has the form

$$\frac{dS(\mathbf{x})}{dt} = -S(\mathbf{x}) + \frac{1}{z}\sum_i S(\mathbf{x}+\mathbf{e}_i) = \Delta S(\mathbf{x})\,, \tag{7.5}$$

where $\mathbf{e_i}$ are the unit vectors of the lattice and Δ denotes the discrete Laplacian operator

$$\Delta F(\mathbf{x}) \equiv -F(\mathbf{x}) + \frac{1}{z}\sum_i F(\mathbf{x}+\mathbf{e}_i)\,.$$

This rate equation shows that the mean spin undergoes a continuous-time random walk on the lattice. As a result, the mean magnetization, $m \equiv \sum_{\mathbf{x}} S(\mathbf{x})/N$ is conserved, as follows by summing Eq. (7.5) over all sites. A subtle aspect of this conservation law is that while the magnetization of a specific system *does* change in a single update event by construction, the average over all sites and over all trajectories of the dynamics *is* conserved. The consequence of this conservation law is profound. Consider a finite system with an initial fraction ρ of Democrats and $1-\rho$ of Republicans; equivalently, the initial magnetization $m_0 = 2\rho - 1$. Ultimately, this system will reach consensus by voter model dynamics — Democrat consensus with probability $E(\rho)$ and Republican consensus with probability $1 - E(\rho)$. The magnetization of this final state is $m_\infty = E(\rho)\times 1 + (1-E(\rho))\times(-1) = 2E(\rho) - 1$. Using magnetization conservation, we obtain a basic conclusion about the voter model: because $m_\infty = m_0$, the "exit probability" is $E(\rho) = \rho$.

Discrete Diffusion Equation and Bessel Functions

For a continuous-time nearest-neighbor lattice random walk, the master equation for the probability that the particle is at site n at time t has the generic form:

$$\dot{P}_n = \frac{\gamma}{2}(P_{n-1} + P_{n+1}) - P_n. \tag{7.6}$$

The random walk corresponds to $\gamma = 1$ in which the total probability is conserved. Here we consider general values of γ because this case arises in the equations of motion for correlation functions in the kinetic Ising model. For simplicity, suppose that the random walk is initially at site $n = 0$. To solve this equation, we introduce the Fourier transform $P(k,t) = \sum_n P_n(t)e^{ikn}$ and find that the Fourier transform satisfies $\frac{dP(k,t)}{dt} = [\frac{1}{2}\gamma(e^{ik} + e^{-ik}) - 1]P(k,t)$. For the initial condition $P(k, t=0) = 1$, the solution is simply $P(k,t) = \exp[\gamma t\cos k - t]$. Now we use the generating function representation of the Bessel function,

$$\exp(z\cos k) = \sum_{n=-\infty}^{\infty} e^{ikn} I_n(z).$$

Expanding the generating function in a power series in γt, we obtain the final result

$$P_n(t) = I_n(\gamma t)e^{-t}. \tag{7.7}$$

In the long-time limit, we use the asymptotics of the Bessel function $I_n(t) \sim (2\pi t)^{-1/2}\, e^t$, to give the asymptotic behavior

$$P_n(t) \sim \frac{1}{\sqrt{2\pi\gamma t}}\, e^{-(1-\gamma)t}.$$

Let's now solve the rate equation (7.5) explicitly for the mean spin at $\mathbf{x}$ for the initial condition $S(x, t = 0) = \delta_{x,0}$; that is, a single Democrat in a background population of undecided voters. In one dimension, the rate equation is

$$\frac{dS(x)}{dt} = -S(x) + \frac{1}{2}\left[S(x-1) + S(x+1)\right]. \tag{7.8}$$

Using the results from the above highlight on the Bessel function solution to this type of master equation, we simply have

$$S(x,t) = I_x(t)\, e^{-t} \sim \frac{1}{\sqrt{2\pi t}} \qquad \text{as } t\to\infty. \tag{7.9}$$

Exactly the same approach works in higher dimensions. Now we introduce the multidimensional Fourier transform $P(k_1, k_2, \ldots, t) = \sum_{x_1,x_2,\ldots} P_{x_1,x_2,\ldots}(t) e^{ik_1x_1} e^{ik_2x_2} \ldots$ and find that the Fourier transform in each coordinate direction factorizes. For the initial condition of one Democrat at the origin in a sea of undecided voters, the mean spin is then given by

$$S(\mathbf{x}, t) = \prod_{i=1}^{d} I_{x_i}(t)\, e^{-dt} \sim \frac{1}{(2\pi t)^{d/2}} \,. \tag{7.10}$$

Thus the fate of a single voter is to quickly relax to the average opinion of the rest of the population — namely everyone is undecided, on average.

While the above result is exact, it provides no information about how consensus is actually achieved in the voter model. What we need is a quantity that tells us the extent to which two distant voters agree. Such a measure is provided by the two-point correlation function, $S(\mathbf{x}, \mathbf{y}) \equiv \langle s(\mathbf{x}) s(\mathbf{y}) \rangle$. Proceeding in close analogy with Eq. (7.3) the two-point function evolves as

$$s(\mathbf{x}, t+\Delta t) s(\mathbf{y}, t+\Delta t) = \begin{cases} s(\mathbf{x}, t) s(\mathbf{y}, t) & \text{with probability } 1 - [w(s(\mathbf{x})) + [w(s(\mathbf{y}))]\Delta t, \\ -s(\mathbf{x}, t) s(\mathbf{y}, t) & \text{with probability } [w(s(\mathbf{x})) + w(s(\mathbf{y}))]\Delta t. \end{cases} \tag{7.11}$$

Thus $(\mathbf{x})s(\mathbf{y}$ changes by $-2(\mathbf{x})s(\mathbf{y}$ if either of the voters at $\mathbf{x}$ or $\mathbf{y}$ changes state with respective rates $w(s(\mathbf{x}))$ and $w(s(\mathbf{y}))$, so that $S(\mathbf{x}, \mathbf{y})$ evolves according to

$$\frac{dS(\mathbf{x}, \mathbf{y})}{dt} = -2 \left\langle s(\mathbf{x}) s(\mathbf{y}) \big[w(s(\mathbf{x})) + w(s(\mathbf{y})) \big] \right\rangle .$$

On a hypercubic lattice, the explicit form of this rate equation is

$$\frac{dS(\mathbf{x}, \mathbf{y})}{dt} = -2S(\mathbf{x}, \mathbf{y}) + \sum_i \frac{1}{z} \big[S(\mathbf{x} + \mathbf{e}_i, \mathbf{y}) + S(\mathbf{x}, \mathbf{y} + \mathbf{e}_i) \big] . \tag{7.12}$$

In what follows, we discuss spatially homogeneous and isotropic systems in which the correlation function depends only on the distance $r = |\mathbf{x} - \mathbf{y}|$ between two voters at $\mathbf{x}$ and $\mathbf{y}$; thus $G(r) \equiv S(\mathbf{x}, \mathbf{y})$. Then the last two terms on the right-hand side of (7.12) are identical and this equation reduces to (7.5) apart from an overall factor of 2. It is now convenient to consider the continuum limit, for which Eq. (7.12) reduces to the diffusion equation

$$\frac{\partial G}{\partial t} = D \nabla^2 G, \tag{7.13}$$

with D is the diffusion coefficient associated with the continuum limit of (7.12). For the undecided initial state in which each voter is independently a Democrat or a Republican with equal probability, the initial condition is $G(r, t = 0) = 0$ for $r > 0$. On the other hand, each voter is perfectly correlated with itself, that is $S(\mathbf{x}, \mathbf{x}) = 1$. In the continuum limit, we must impose a lower cutoff a in the argument of the correlation function, so that the statement of perfect self correlation becomes $G(a, t) = 1$.

To understand physically how the correlation function evolves, it is expedient to work with $c \equiv 1 - G$; c also satisfies the diffusion equation, but now with the initial condition $c(r > a, t = 0) = 1$, and the boundary condition $c(r = a, t) = 0$; that is, the absorbing point at the origin is replaced by a small absorbing sphere of non-zero radius a. One should think of a as playing the role of the lattice spacing; a non-zero radius is needed so that a diffusing particle can actually hit the sphere. Physically, then, we study how an initially constant density profile evolves in the presence of a small absorbing sphere at the origin. The exact solution for this concentration profile can be easily obtained in the Laplace domain. Laplace transforming the diffusion equation gives $sc - 1 = D\nabla^2 c$; the inhomogeneous term arises from the constant-density initial condition. A particular solution to the inhomogeneous equation is simply $c = 1/s$, and the homogeneous equation

$$c'' + \frac{d-1}{r}\, c' - \frac{s}{D}\, d = 0$$

has the general solution $c = Ar^\nu I_\nu(r\sqrt{s/D}) + Br^\nu K_\nu(r\sqrt{s/D})$, where I_ν and K_ν are the modified Bessel functions of order ν, with $\nu = (2-d)/2$. Since the concentration is finite as $r \to \infty$, the term with I_ν must

be rejected. Then matching to the boundary condition $c = 0$ at $r = a$ gives

$$c(r,s) = \frac{1}{s}\left[1 - \left(\frac{r}{a}\right)^{\nu} \frac{K_\nu(r\sqrt{s/D})}{K_\nu(a\sqrt{s/D})}\right]. \tag{7.14}$$

For spatial dimension $d > 2$, corresponding to $\nu < 0$, we use $K_\nu = K_{-\nu}$ and the small-argument form $K_\nu(x) \propto (2/x)^\nu$ to give the leading small-s behavior

$$c(r, s \to 0) = \frac{1}{s}\left[1 - \left(\frac{a}{r}\right)^{d-2}\right].$$

Thus in the time domain, the concentration profile approaches the static electrostatic solution, $c(r) = 1 - (a/r)^{d-2}$! A steady state is achieved because there is a non-zero probability that a diffusing particle never hits the absorbing sphere. This is the phenomenon of *transience* that was discussed in Sec. (2.4). The depletion of the concentration near the sphere is sufficiently slow that it is replenished by re-supply from more distant particles. In terms of the voter model, the two-particle correlation function asymptotically becomes $G(r) \to (a/r)^{d-2}$ for $d > 2$. Thus the influence of one voter on a distant neighbor decays as a power law in their separation.

Now let's study the case $d \leq 2$ ($\nu \geq 0$). Here a diffusing particle eventually hits the sphere; this is the property of *recurrence* (see again Sec. 2.4) that leads to a growing depletion zone about the sphere. While the time dependence of c can be obtained by inverting the Laplace transform in Eq. (7.14), it is much simpler to apply the quasi-static approximation as first outlined in Sec. 2.5. From the results given in that section [Eqs. (2.52a) and (2.52b)], the two-spin correlation function for $r > a$ has the asymptotic behavior for general spatial dimensions:

$$G(r,t) \sim \begin{cases} 1 - \left(\dfrac{r}{\sqrt{Dt}}\right)^{2-d} & d < 2 \text{ and } 0 < r < \sqrt{Dt}; \\[2ex] 1 - \dfrac{\ln(r/a)}{\ln(\sqrt{Dt}/a)} & d = 2 \text{ and } a < r < \sqrt{Dt}; \\[2ex] \left(\dfrac{a}{r}\right)^{d-2} & d > 2 \text{ and } a < r. \end{cases} \tag{7.15}$$

An important feature for $d \leq 2$ is that the correlation function at fixed r approaches 1—distant spins gradually become more strongly correlated. This feature is a manifestation of *coarsening* in which the voters organize into a mosaic of single-opinion enclaves whose characteristic size increases with time. As we shall discuss in more detail in chapter 8, coarsening typifies many types of phase-ordering kinetics. On the other hand, for $d > 2$ the voter model approaches a steady state and there is no coarsening in the spatial arrangement of the voters if the population is infinite.

There are two important consequences for the voter model that can be deduced from the behavior of the correlation function. The first is that we can immediately determine the time dependence of the density of "interfaces", namely, the fraction n of neighboring voters of the opposite opinion. As we shall use extensively later in this chapter, it is helpful to represent an interface as an effective particle that occupies the bond between two neighboring voters of the opposite opinion. This effective particle, or domain wall, provides the right way to characterize the departure of system from consensus. For nearest-neighbor sites $\mathbf{x}$ and $\mathbf{y}$, we relate the correlation function to the domain wall density by

$$\begin{aligned} G(\mathbf{x},\mathbf{y}) = \langle s(\mathbf{x})s(\mathbf{y})\rangle &= [\text{prob}(++) + \text{prob}(--)] - [\text{prob}(+-) + \text{prob}(-+)] \\ &= 1 - n \quad - \quad n = 1 - 2n. \end{aligned} \tag{7.16}$$

Thus the density of interfaces is related to the near-neighbor correlation function via $n = (1 - G(\mathbf{x},\mathbf{y}))/2$. Using our result (7.15) for the correlation function, the time dependence of the interfacial density is then

$$n(t) \sim \begin{cases} t^{d/2-1} & d < 2, \\ 1/\ln t & d = 2, \\ \mathcal{O}(1) & d > 2. \end{cases} \tag{7.17}$$

When $d \leq 2$, the probability of having two voters with opposite opinions asymptotically vanishes and the system develops a coarsening mosaic of single-opinion domains (Fig. 7.1). At the marginal dimension of $d = 2$ the coarsening process is very slow and the density of interfaces asymptotically vanishes as $1/\ln t$. In higher dimensions, an infinite system reaches a dynamic frustrated state where voters of opposite opinion coexist and continually evolve such that the mean density of each type of voter remains fixed.

The second basic consequence that follows from the correlation function is the time T_N to reach consensus for a finite system of N voters. For this estimate of the consensus time, we use the fact that the influence of any voter spreads diffusively through the system. Thus starting with some initial state, the influence range of one voter is of the order of $\sqrt{Dt}$. We then define consensus to occur when the total amount of correlation within a distance of $\sqrt{Dt}$ of a particular voter equals the total number of voters N. The consensus criterion therefore becomes

$$\int^{\sqrt{Dt}} G(r)\, r^{d-1}\, dr = N. \tag{7.18}$$

The lower limit can be set to 0 for $d = 1$ and should be set to a for $d > 1$. Substituting the expressions for the correlation function given in Eq. (7.15) into this integral, the time dependence can be extracted merely by scaling and we find the asymptotic behavior

$$T_N \propto \begin{cases} N^{2/d} & d < 2; \\ N \ln N & d = 2; \\ N & d > 2. \end{cases}$$

Thus as the dimension decreases below 2, consensus takes a progressively longer to achieve. This feature reflects the increasing difficulty in transmitting information when the dimensionality decreases.

This last part of the section hanging and incomplete. Let us now derive the exact solution for the correlation function without using the continuum approximation. This solution is nothing more than the lattice Green's function for the diffusion equation. It is convenient to rescale the time variable by 2, $\tau = 2t$, so that the correlation function satisfies precisely the same equation of motion as the average magnetization

$$\frac{d}{d\tau} G(\mathbf{x}) = -G(\mathbf{x}) + \frac{1}{z} \sum_i G(\mathbf{x} + \mathbf{e}_i). \tag{7.19}$$

We consider the uncorrelated initial condition $G(\mathbf{x}, 0) = \delta(\mathbf{x})$ and the boundary condition is $G(\mathbf{0}) = 1$. The evolution equation and the initial conditions are as for the autocorrelation function where the solution is $I_{\mathbf{m}}(\tau) e^{-d\,\tau}$. Since the equation is linear, every linear combination of these "building-blocks" is also a solution. Therefore, we consider the linear combination

$$G(\mathbf{x}, \tau) = I_{\mathbf{x}}(\tau) e^{-d\,\tau} + \int_0^\tau d\tau' J(\tau - \tau') I_{\mathbf{x}}(\tau') e^{-d\,\tau'}. \tag{7.20}$$

The kernel of the integral is identified as a source with strength $\delta(\tau) + J(\tau)$. This source is fixed by the boundary condition:

$$1 = \left[I_0(\tau) e^{-\tau}\right]^d + \int_0^\tau d\tau'\, J(\tau - \tau') [I_0(\tau') e^{-\tau'}]^d. \tag{7.21}$$

We are interested in the asymptotic behavior of the correlation function. This requires the $\tau \to \infty$ behavior of the source term. Thus, we introduce the Laplace transform $\hat{J}(s) = \int_0^\infty d\tau\, e^{-s\tau} J(\tau)$. Exploiting the convolution structure of the integral yields

$$\hat{J}(s) = [s\hat{I}(s)]^{-1} - 1 \qquad \text{with} \qquad \hat{I}(s) = \int_0^\infty d\tau\, e^{-s\tau} [I_0(\tau) e^{-\tau}]^d. \tag{7.22}$$

Using the integral representation of the Bessel function, $I_0(\tau) = \int_0^{2\pi} \frac{dq}{2\pi} e^{\tau \cos q}$, the latter transform is expressed as an integral

$$\hat{I}(s) = \int_0^{2\pi} \frac{d\mathbf{q}}{(2\pi)^d}\, \frac{1}{s + \sum_{i=1}^d (1 - \cos q_i)}. \tag{7.23}$$

The $\tau \to \infty$ asymptotic behavior of the source and the correlation function is ultimately related to the $s \to 0$ asymptotic behavior of this integral. The integral diverges, $\hat{I}(s) \sim s^{d/2-1}$, when $d < 2$, but it remains finite when $d > 2$. The leading $s \to 0$ behavior of the Laplace transform is therefore

$$\hat{J}(s) \sim \begin{cases} s^{-d/2} & d < 2, \\ s^{-1} \ln s^{-1} & d = 2, \\ s^{-1} & d > 2. \end{cases} \tag{7.24}$$

7.2 Glauber Model in One Dimension

In the Ising model, a regular lattice is populated by 2-state spins that may take one of two values: $s(\mathbf{x}) = \pm 1$. Pairs of nearest-neighbor spins experience a ferromagnetic interaction that favors their alignment. The Hamiltonian of the system is

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j, \tag{7.25}$$

where the sum is over nearest neighbors (i, j) on the lattice. Every parallel pair of neighboring spins contributes $-J$ to the energy and every antiparallel pair contributes $+J$. When the coupling constant is positive, the interaction favors ferromagnetic order. The main feature of the Ising model is that ferromagnetism appears spontaneously in the absence of any driving field when the temperature T is less than a critical temperature T_c and the spatial dimension $d > 1$. Above T_c, the spatial arrangement of spins is spatially disordered, with equal numbers of spins in the states $+1$ and -1. Consequently, the magnetization is zero and spatial correlations between spins decay exponentially with their separation. Below T_c, the magnetization is non-zero and distant spins are strongly correlated. All thermodynamic properties of the Ising model can be obtained from the partition function $Z = \sum \exp(-\beta \mathcal{H})$, where the sum is over all spin configurations of the system, with $\beta = 1/kT$ and k is the Boltzmann constant.

While equilibrium properties of the Ising model follow from the partition function, its non-equilibrium properties depend on the nature of the spin dynamics. There is considerable freedom in formulating this dynamics that is dictated by physical considerations. For example, the spins may change one at a time or in correlated blocks. More fundamentally, the dynamics may or may not conserve the magnetization. The role of a conservation law depends on whether the Ising model is being used to describe alloy systems, where the magnetization (related to the composition of the material) is necessarily conserved, or spin systems, where the magnetization does not have to be conserved. This lack of uniqueness of dynamical rules is generic in non-equilibrium statistical physics and it part of the reason why there do not exist universal principles, such as free energy minimization in equilibrium statistical mechanics, that definitive prescribe how a non-equilibrium spin system evolves.

Spin evolution

We now discuss a simple version of the kinetic Ising model — first introduced by Glauber in 1963 — with non-conservative single-spin-flip dynamics that allows one to extend the Ising model to non-equilibrium processes. We first focus on the exactly-soluble one-dimensional system, and later in this chapter we will study the Ising-Glauber model in higher dimensions, and well as different types of spin dynamics, including conservative Kawasaki spin-exchange dynamics, and cluster dynamics, in which correlated blocks of spins flip simultaneously. In the Glauber model, spins are selected one at a time in random order and each changes at a rate that depends on the change in the energy of the system as a result of this update. Because only single spins can change sign in an update, $s_j \to -s_j$, where s_j is the spin value at site j, the magnetization is generally *not* conserved.

There are three types of transitions when a single spin flips: energy raising, energy lowering, and energy neutral transitions (Fig. 7.2). Energy raising events occur when a spin is aligned with a majority of its neighbors and *vice versa* for energy lowering events. Energy conserving events occur when the magnetization of the neighbors is zero. The basic principle to fix the rates of these of events is the *detailed balance condition*. Mathematically, this condition is:

$$P(\{s\})w(s \to s'_j) = P(\{s'_j\})\, w(s'_j \to s). \tag{7.26}$$

Here $\{s\}$ denotes the state of all the spins in the system, $\{s'_j\}$ denotes the state derived from $\{s\}$ in which the spin at i is flipped, and $w(s \to s'_j)$ denotes the transition rate from $\{s\}$ to $\{s'_j\}$.

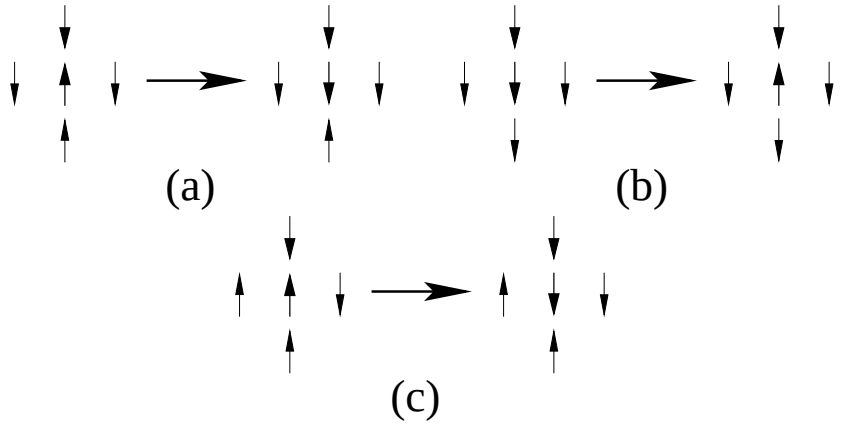


Figure 7.2: (a) Energy lowering, (b) energy raising, and (c) energy conserving spin-flip events on the square lattice.

The detailed balance condition is merely a statement of current conservation. In the abstract space of all 2^N possible spin states of a system of N spins, Glauber dynamics connects states which differ by the flipping of a single spin. When detailed balance holds, the probability currents from state $\{s\}$ to $\{s'_j\}$ and from $\{s'_j\}$ to $\{s\}$ (the left and right sides of Eq. (7.26)) are equal so there is no net probability current across any link in this state space. If $P(\{s\})$ are the equilibrium Boltzmann weights, then the transition rates defined by Eq. (7.26) ensure that any initial spin state will eventually relax to the equilibrium thermodynamic equilibrium state for any non-zero temperature. Thus dynamics that satisfy detailed balance are required if one seeks to understand how equilibrium is approached when a system is prepared in an out-of-equilibrium state.

In one dimension, the detailed balance condition is sufficient to actually fix the flip rates. Following Glauber, we assume that the flip rate of the $j^{\rm th}$ spin depends on the neighbors with which there is a direct interaction, namely, s_j and $s_{j\pm1}$. For an isotropic system, the rate should have left/right symmetry (invariance under the interchange $i+1 \leftrightarrow i-1$) and up/down symmetry (invariance under the reversal of all spins)[1]. For a homogeneous one-dimensional system, these conditions constrain the rate to have the form $w(s \to s'_j) = A + Bs_j(s_{j-1} + s_{j+1})$. This flip rate is simply the energy of the $i^{\rm th}$ spin up to an additive constant. We now write this flip rate in the following suggestive form

$$w(s \to s'_j) = \frac{\alpha}{2}\left[1 - \frac{\gamma}{2}s_j(s_{j-1} + s_{j+1})\right] = \begin{cases} \frac{\alpha}{2}(1-\gamma) & \text{for spin state } \uparrow\uparrow\uparrow \text{ or } \downarrow\downarrow\downarrow; \\ \frac{\alpha}{2} & \text{for spin state } \uparrow\uparrow\downarrow \text{ or } \downarrow\downarrow\uparrow; \\ \frac{\alpha}{2}(1+\gamma) & \text{for spin state } \uparrow\downarrow\uparrow \text{ or } \downarrow\uparrow\downarrow\,. \end{cases} \tag{7.27}$$

When the two neighbors are antiparallel (no local field), the flip rate is simply a constant that we take to be 1/2 ($\alpha = 1$) without loss of generality. For $\gamma > 0$, the flip rate favors aligning s_j with its neighbors and *vice versa* for $\gamma < 0$.

We now fix γ by applying detailed balance:

$$\frac{w(s \to s'_j)}{w(s'_j \to s)} = \frac{1 - \frac{\gamma}{2}s_j(s_{j-1} + s_{j+1})}{1 + \frac{\gamma}{2}s_j(s_{j-1} + s_{j+1})} = \frac{P(\{s'_j\})}{P(\{s\})} = \frac{e^{+\beta J\epsilon_j}}{e^{-\beta J\epsilon_j}}, \tag{7.28}$$

with $\epsilon_j \equiv -s_j(s_{j-1} + s_{j+1})$. We simplify the last quantity by exploiting the ±1 algebra of Ising spins to write

$$\frac{e^{+\beta J\epsilon_j}}{e^{-\beta J\epsilon_j}} = \frac{\cosh\beta J\epsilon_j + \sinh\beta J\epsilon_j}{\cosh(-\beta J\epsilon_j) + \sinh(-\beta J\epsilon_j)} = \frac{1 + \tanh(2\beta J\frac{\epsilon_j}{2})}{1 - \tanh(2\beta J\frac{\epsilon_j}{2})} = \frac{1 + \frac{1}{2}\epsilon_j \tanh 2\beta J}{1 - \frac{1}{2}\epsilon_j \tanh 2\beta J},$$

[1] Actually the most general rate that satisfies the constraints of locality within the interaction range, symmetry, and isotropy is $w_(s_j) = (1/2)(1 + \delta s_{j-1}s_{j+1})\left[1 - (\gamma/2)s_j(s_{j-1} + s_{j+1})\right]$

where in the last step we use the fact that $\tanh ax = a\tanh x$ for $a = 0, \pm 1$. Comparing with Eq. (7.27), we deduce that $\gamma = \tanh 2\beta J$. Thus the flip rate is

$$w(s_j) = \frac{1}{2}\left[1 - \frac{1}{2}\tanh 2\beta J\, s_j(s_{j-1} + s_{j+1})\right]. \tag{7.29}$$

For $T \to \infty$, $\gamma \to 0$ and all three types of spin-flip events shown in Eq. (7.27) are equiprobable. Conversely, for $T \to 0$, $\gamma \to 1$, and energy raising spin-flip events are prohibited.

The probability distribution $P(\{\mathbf{s}\}, t)$ that the system has the microscopic spin configuration $\mathbf{s}$ at time t satisfies the same master equation (7.2) as the voter model. Consequently, the equation of motion for the low-order correlation functions are:

$$\frac{dS_j}{dt} = -2\big\langle s_j w(s_j)\big\rangle, \tag{7.30a}$$

$$\frac{dS_{i,j}}{dt} = -2\big\langle s_i s_j[w(s_i) + w(s_j)]\big\rangle, \tag{7.30b}$$

where the subscripts i and j denote the i^{th} and j^{th} site of a one-dimensional lattice.

Using the transition rates given in (7.29) and the identity $s_j^2 = 1$, the rate equation for the average spin S_j is

$$\frac{dS_j}{dt} = -S_j + \frac{\gamma}{2}\,(S_{j-1} + S_{j+1})\,. \tag{7.31}$$

With the initial condition $S_j(0) = \delta_{j,0}$, the solution is (see the highlight on page 117 on the Bessel function solution to discrete diffusion)

$$S_j(t) = I_j(\gamma t)e^{-t}. \tag{7.32}$$

The new feature compared to the corresponding voter model solution is the presence of the temperature-dependent factor γ. Now the average spin at any site decays as $S_j(t) \sim (2\pi\gamma t)^{-1/2}e^{-(1-\gamma)t}$. For $T > 0$, the decay is exponential in time, $S_j \sim e^{-t/\tau}$, with relaxation time $\tau = (1-\gamma)^{-1}$, while for $T = 0$ the decay is algebraic in time, $S_j \simeq (2\pi t)^{-1/2}$. The magnetization $m = N^{-1}\sum_j S_j$ satisfies $\frac{dm}{dt} = -(1-\gamma)m$, so that m decays exponentially with time at any positive temperature,

$$m(t) = m(0)e^{-(1-\gamma)t}\,, \tag{7.33}$$

and is conserved only at zero temperature, just as in the voter model. The Ising-Glauber in one dimension model illustrates critical slowing down — slower relaxation at the critical point ($T = 0$ in one dimension) than for $T > 0$.

The mean spin can also be directly solved for a general initial condition, $S_j(t = 0) = \sigma_j$, with σ_j an arbitrary function between $+1$ and -1. Then the Fourier transform of the initial condition is $s_k(t = 0) = \sum_n \sigma_n e^{ikn}$. Using this result, the Fourier transform of the solution to the equation of motion (7.31) is

$$S_k(t) = S_k(t=0)e^{(\gamma\cos k - 1)t} = \sum_m e^{ikm}\sigma_m \sum_n I_n(\gamma t)e^{ikn}\, e^{-t}\,.$$

Now define $\ell = m + n$ to recast the exponential factors as a single sum to facilitate taking the inverse Fourier transform:

$$S_k(t) = \sum_\ell e^{ik\ell}\sum_m \sigma_m\, I_{\ell-m}(\gamma t)\, e^{-t}\,.$$

From the expression above we may simply read off the solution as the coefficient of $e^{ik\ell}$:

$$S_\ell = \sum_m \sigma_m\, I_{\ell-m}(\gamma t)\, e^{-t}\,. \tag{7.34}$$

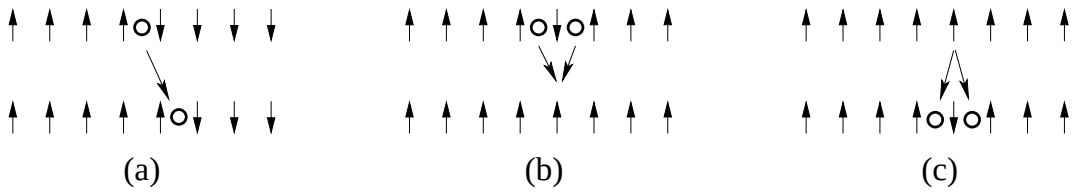


Figure 7.3: Mapping between states of the Ising or the voter models in one dimension and domain wall particles between neighboring pairs of antiparallel spins. Shown are the equivalences between: (a) an energy conserving move and diffusion of a domain wall, (b) energy lowering moves and annihilation of two domain walls, and (c) energy raising moves and creation of a pair of domain walls.

Let's now study the pair correlation function, $S_{i,j} = \langle s_i s_j \rangle$. As a preliminary, we highlight a geometrical equivalence between the kinetic Ising model and diffusion-limited reactions. As given by Eq. (7.16), there is a one-to-one mapping between a spin configuration and an arrangement of domain wall quasi particles. Two neighboring antiparallel spins are equivalent to a domain wall that is halfway between the two spins, while two neighboring parallel spins has no intervening domain wall (Fig. 7.3). Energy raising spin flips are equivalent to creating a nearest-neighbor pair of domain walls, while energy lowering moves correspond to annihilation of two neighboring walls. Energy conserving flips correspond to the hopping of a domain wall between neighboring sites. At $T = 0$, where domain wall creation is forbidden, Ising-Glauber kinetics is then equivalent to *irreversible diffusion-controlled annihilation*, $A + A \to 0$ that we will discuss in more detail in chapter 9.

We focus on translationally invariant systems where the correlation function depends only the separation of the two spins, $G_k \equiv S_{i,i+k}$. The master equation (7.30b) becomes

$$\frac{dG_k}{dt} = -2G_k(t) + \gamma\,(G_{k-1} + G_{k+1}) \tag{7.35}$$

for $k > 0$. This equation needs to be supplemented by the boundary condition $G_0(t) = 1$. Thus the pair correlation function evolves in nearly the same way as the mean spin. However, because of the existence of the fixed boundary condition at the origin, the master equation also admits an exponential equilibrium solution. This solution is determined by assuming that $G_k(\infty) \propto \eta^k$ and substituting this form into Eq. (7.35) with the left-hand side set to zero. These steps lead to the following condition for η: $2 = \gamma(\eta + \eta^{-1})$, whose solution is $\eta = [1 - \sqrt{1-\gamma^2}]/\gamma = \tanh \beta J$. The equilibrium pair correlation function therefore decays exponentially in the distance between the two spins,

$$G_k(\infty) = e^{-k/\xi}\,, \tag{7.36}$$

with correlation length $\xi^{-1} = \ln(\coth \beta J)$. This result coincides with the correlation function obtained directly from thermodynamics. As expected, the correlation length ξ diverges as $T \to 0$, indicative of a phase transition, and ξ vanishes at infinite temperature.

To solve the time dependence of the correlation function with the prescribed initial and boundary conditions, we use the fact that master equation for the correlation function has the same form as that for the mean spin, apart from an overall factor of 2. Thus the general solution will be built from components of the same form as (7.34), with the replacement of $\gamma \to 2\gamma$. We now need to determine the appropriate linear combination of these component solutions that simultaneously satisfy the initial condition $G_k(t = 0)$ and the boundary condition $G_0 = 1$. One piece of the full solution is just the equilibrium correlation function $G_k(\infty) = \eta^{|k|}$. To this we add the general homogeneous solution that satisfies the prescribed constraints. Pictorially, the appropriate initial condition for the homogeneous solution consists of an arbitrary odd function plus an antisymmetric piece that cancels the equilibrium solution for $k > 0$ (Fig. 7.4). The antisymmetry of these pieces ensure that $G_0 = 1$ and that the prescribed initial condition is satisfied for $k > 0$.

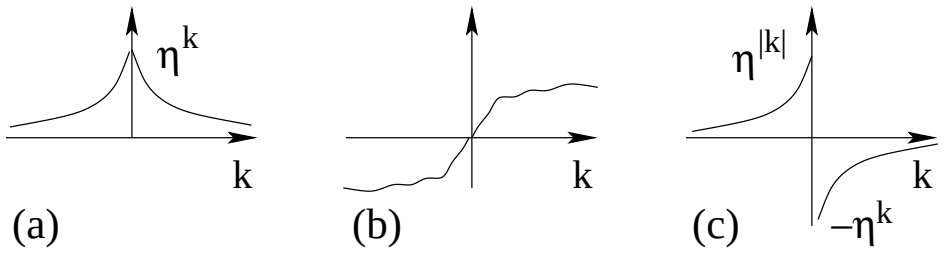


Figure 7.4: (a) Equilibrium correlation function and (b) an arbitrary antisymmetric initial condition. To find $G_k(t)$ for $k > 0$, we superpose the solutions for the three initial conditions shown. The influence of the initial condition (c) cancels that of (a) in the region $k > 0$ so that only (b) remains—an arbitrary initial condition that vanishes at $k = 0$.

The general solution for $k > 0$ therefore is:

$$\begin{aligned} G_k(t) &= \eta^k + e^{-2t} \sum_{\ell=-\infty}^{\infty} G_\ell(0) I_{k-\ell}(2\gamma t) \\ &= \eta^k + e^{-2t} \sum_{\ell=1}^{\infty} [G_\ell(0) - \eta^\ell] I_{k-\ell}(2\gamma t) + e^{-2t} \sum_{\ell=-1}^{-\infty} [G_\ell(0) + \eta^{|\ell|}] I_{k-\ell}(2\gamma t) \\ &= \eta^k + e^{-2t} \sum_{\ell=1}^{\infty} [G_\ell(0) - \eta^\ell][I_{k-\ell}(2\gamma t) - I_{k+\ell}(2\gamma t)]. \end{aligned} \tag{7.37}$$

We restrict ourselves to the case of $T = 0$, where two special cases lead to nice results:

1. Antiferromagnetic initial state, $G_k(0) = (-1)^k$. In this case, every site of the dual lattice is initially occupied by domain wall particle. For this initial state, the nearest-neighbor correlation function in Eq. (7.37) reduces to

$$G_1(t) = 1 - 2e^{-2t} \sum_{j\,\text{odd}} [I_{1-j}(2t) - I_{1+j}(2t)] = 1 - 2e^{-2t}\, I_0(2t),$$

   where we have used $I_n = I_{-n}$.

2. Random initial state, $G_k(0) = m_0^2$, where m_0 is the initial magnetization. Then the nearest-neighbor correlation function is

$$G_1(t) = 1 + e^{-2t}(m_0^2 - 1) \sum_j [I_{1-j}(2t) - I_{1+j}(2t)] = 1 - 2e^{-2t}\,(m_0^2 - 1)\,[I_0(2t) + I_1(2t)]\,.$$

From these two solutions, the domain wall densities are

$$\rho(t) = \frac{1 - G_1}{2} = \begin{cases} I_0(2t)\, e^{-2t} \sim \dfrac{1}{\sqrt{4\pi t}} & \text{antiferromagnetic,} \\ \dfrac{1 - m_0^2}{2}\,[I_0(2t) + I_1(2t)]\; e^{-2t} \sim \dfrac{1 - m_0^2}{\sqrt{4\pi t}} & \text{uncorrelated.} \end{cases} \tag{7.38}$$

If the initial magnetization $m_0 = 0$ for the random initial condition, then the asymptotic domain wall density universally vanishes as

$$\rho(t) \sim (4\pi t)^{-1/2}, \tag{7.39}$$

independent of the initial domain wall density! Because the number of domain walls decrease with time, their separation correspondingly increases. The system therefore coarsens, as domains of parallel spins grow with the diffusive length scale $t^{1/2}$. A final important point is that Eq. (7.38) also represents the exact solution for diffusion-limited annihilation $A + A \to 0$! We will return to this reaction in chapter 9.

Domain length distribution

In the previous section, we obtained the density of domain walls or alternatively, the average domain length. Now we ask the more fundamental question: what is the distribution of domain lengths in a one-dimensional system of length L? Let P_k be the probability for a domain of length k, namely, a string of k consecutive aligned spins that is flanked by oppositely-oriented spins at each end. To have a system-size independent quantity, we define this probability per unit length.

Partial information about this distribution follows from basic physical considerations. For example, the domain wall density ρ, which scales as $t^{-1/2}$, is given by $\sum_k P_k$, while the domain length distribution obeys the normalization condition $\sum_k kP_k = 1$. Furthermore, from the diffusive nature of the evolution, the only physical length scale grows as $t^{1/2}$. These facts suggest that the domain length distribution has the scaling form

$$P_k(t) \simeq t^{-1}\Phi(kt^{-1/2}). \tag{7.40}$$

The prefactor ensures that the mean domain length (per unit length) equals 1, *i.e.*, $\int x\Phi(x)dx = 1$, while the asymptotic decay of the total density (7.39) gives the condition $\int \Phi(x)dx = (4\pi)^{-1/2} \equiv C$.

We can also infer the short-distance tail of the scaling function $\Phi(x)$ from the long-time decay of the domain density. Consider the role of the shortest possible domain of length 1 in the rate equation for the domain density ρ. When a domain that consists of a single spin flips, three domains merge into a single larger domain, as illustrated below:

$$\cdots\downarrow\underbrace{\uparrow\cdots\uparrow\uparrow}\underline{\downarrow}\underbrace{\uparrow\uparrow\cdots\uparrow}\downarrow\cdots \xrightarrow{1} \cdots\downarrow\underbrace{\uparrow\cdots\uparrow\uparrow\underline{\uparrow}\uparrow\uparrow\cdots\uparrow}\downarrow\cdots .$$

Since such events, in which two domains disappear, occur with a unit rate, the domain density decays as

$$\frac{d\rho}{dt} = -2P_1, \tag{7.41}$$

and using Eq. (7.39), we obtain $P_1 \sim \frac{C}{4}\,t^{-3/2}$. On the other hand, expanding Φ in a Taylor series gives $P_1 \cong \Phi(0)t^{-1} + \Phi'(0)t^{-3/2} + \cdots$. Comparing these two results, we deduce that $\Phi(0) = 0$ and $\Phi'(0) = \frac{C}{4}$. Therefore the scaling function vanishes linearly in the small-argument limit:

$$\Phi(x) \sim \frac{C}{4}\,x, \qquad \text{as} \qquad x \to 0. \tag{7.42}$$

This linear decrease in the small-size tail of the probability distribution is a generic feature of many one-dimensional interacting many-body systems.

While scaling provides some glimpses about the nature of the length distribution, we are interested in the distribution itself. The exact solution of the distribution is not yet known, and we present an approximate solution based on the *independent interval approximation*. This approximation is based on assuming that the lengths of neigh boring domains are uncorrelated, an assumption makes the calculation of the domain length distribution tractable. This same approximation can be applied to a variety of one-dimensional domain evolution and reaction processes. Under the assumption of uncorrelated domains, their length distribution evolves according to the master equations

$$\frac{dP_k}{dt} = -2P_k + P_{k+1} + P_{k-1}\left(1 - \frac{P_1}{\rho}\right) + \frac{P_1}{\rho^2}\sum_{i+j=k-1} P_i\,P_j - \frac{P_1}{\rho}P_k. \tag{7.43}$$

The first three terms account for length changes due to a domain wall hopping by ± 1 and describe the diffusion of a single domain. The factor $1 - P_1/\rho$ multiplying P_{k-1} ensures that the neighboring domain has length greater than 1, so that the hopping of a domain wall leads to $(k-1, j) \to (k, j-1)$, and not to $(k-1, 1, j) \to (k+j)$. The remaining terms account for changes in the distribution due to mergings. Because any merger requires a domain of length 1, these terms are proportional to P_1. The gain term accounts for the merger of three domains of lengths i, j, and 1, with $i + j + 1 = k$, and the loss term accounts for the merger of a domain of length k with a domain of any length. These master equations apply for any $k \geq 1$, subject to the boundary condition $P_0 = 0$.

It is easy to check that the total density $\rho = \sum_k P_k$ satisfies (7.41) and that $\sum_k k\,\frac{dP_k}{dt} = 0$. Since the typical domain length grows indefinitely with time, we replace the integer k by the continuous variable x, and substitute the scaling form (7.40), as well as $\rho \simeq Ct^{-1/2}$ and $P_1 \simeq \frac{C}{4}t^{-3/2}$ into the master equation to give the integro-differential equation for the scaling function

$$\frac{d^2\Phi}{dx^2} + \frac{1}{2}\frac{d(x\Phi)}{dx} + \frac{1}{4C}\int_0^x \Phi(y)\Phi(x-y)\,dy = 0. \tag{7.44}$$

We now introduce the Laplace transform, $\phi(s) = C^{-1}\int_0^\infty \Phi(x)e^{-sx}dx$, to transform the convolution into a product and reduce this integro-differential equation to the ordinary nonlinear differential equation

$$\frac{d\phi}{ds} = \frac{\phi^2}{2s} + 2s\,\phi - \frac{1}{2s}, \tag{7.45}$$

with the boundary condition $\phi(0) = 1$.

Eq. (7.45) is a Riccati equation and it can be reduced to the second-order linear equation

$$\frac{d^2\psi}{ds^2} + \frac{d\psi}{ds}\left(\frac{1}{s} - 2s\right) - \frac{\psi}{4s^2} = 0.$$

by the standard transformation $\phi(s) = -2s\frac{d\ln\psi(s)}{ds}$. We then eliminate the linear term in this equation by writing $\psi = yv$ and then forcing the term linear in ψ' to be zero. This requirement gives the condition $\ln v' = s - 1/(2s)$, from which we find that the transformation $\phi(s) = 1 - 2s^2 - 2s\frac{d}{ds}\ln y(s)$ reduces the Riccati equation (7.45) to a linear Schrödinger equation

$$\frac{d^2y}{ds^2} + (2 - s^2)y = 0. \tag{7.46}$$

Eq. (7.46) is the parabolic cylinder equation whose solution is a linear combination of the two linearly independent solutions, $y(s) = C_+D_{1/2}(s\sqrt{2}) + C_-D_{1/2}(-s\sqrt{2})$, with $D_\nu(x)$ the parabolic cylinder function of order ν. From the large-s behavior $\phi(s) \simeq (4s)^{-2}$, together with the asymptotics of $D_\nu(s)$, it follows that $C_- = 0$. Therefore the Laplace transform is

$$\phi(s) = 1 - 2s^2 - 2s\frac{d}{ds}\ln D_{1/2}(s\sqrt{2}). \tag{7.47}$$

The constant C_+ can be evaluated explicitly from the normalization condition $\phi'(0) = -C_+^{-1}$ and the properties[2] of $D_\nu(x)$. Using these facts, we find $C_+ = \Gamma(3/4)/\Gamma(1/4) = 0.337989\ldots$; this result should be compared with the exact value $C = (4\pi)^{-1/2} = 0.28209$.

The domain length distribution at large length can also be obtained from the small-s limit of the exact solution (7.47). The large-x tail of $\Phi(x)$ is exponential as follows from the behavior of the Laplace transform near its simple pole at $s = -\lambda$, $\phi(s) \simeq 2\lambda(s+\lambda)^{-1}$. The constant λ is given by the first zero of the parabolic cylinder function, $D_{1/2}(-\lambda\sqrt{2}) = 0$, located at $\lambda \approx 0.5409$. Therefore the domain length distribution asymptotically decays exponentially for large x

$$\Phi(x) \simeq A\exp(-\lambda x), \tag{7.48}$$

with amplitude $A = 2C\lambda$. The approximate value for the decay coefficient λ is larger than the exact value $\zeta(3/2)/\sqrt{16\pi} = 0.368468$.

While the independent interval approximation is not exact, it is very useful. By invoking the this approximation, we are able to write a closed master equation for the evolution of the domain length distribution. The independent interval approximation then yields the main qualitative behavior of the domain length distribution including: (i) the linear small-length limit of the distribution, (ii) the large-length exponential tail, and (iii) correct integrated properties, such as the $t^{-1/2}$ decay of the number of domains. As we shall see in later applications, the independent interval approximation applies to a wide range of coarsening processes.

[2]The following properties are needed $D_\nu(0) = \frac{\sqrt{\pi}2^\nu}{\Gamma(1/2-\nu/2)}$, $D'_\nu(0) = \frac{\sqrt{\pi}2^{\nu+1}}{\Gamma(-\nu/2)}$, and $D_\nu(x) \sim x^\nu\exp(-x^2/4)[1 + \mathcal{O}(x^{-2})]$.

7.3 Glauber Model in Higher Dimensions

Finite spatial dimension

When the spatial dimension d is greater than one, the Ising-Glauber model is no longer solvable. A variety of approximate continuum theories have been constructed to capture the essence of this model, as will be described in the next chapter. Here we focus on the basic properties of an individual-spin description, there is still much that can be learned.

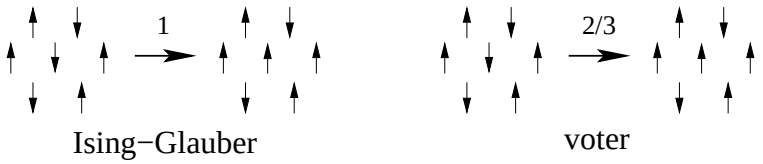


Figure 7.5: Comparison of the rates of an update event in the Ising-Glauber model at zero temperature and in the voter model on the triangular lattice.

First, we address why the voter model is soluble for all d, while the closely related Ising-Glauber model at zero temperature is not. This dichotomy stems from a simple but profound difference between these two models in $d > 1$.[3] Let's determine the transition rates for Glauber dynamics for $d > 1$ by using detailed balance. Following the same steps that lead to Eq. (7.29), we have

$$\frac{w(s \to s_i')}{w(s_i' \to s)} = \frac{P(\{s_i'\})}{P(\{s\})} = \frac{e^{-\beta J s_i \sum s_j}}{e^{+\beta J s_i \sum s_j}} = \frac{1-\tanh(\beta J s_i \sum s_j)}{1+\tanh(\beta J s_i \sum s_j)} = \frac{1 - s_i \tanh \beta J \sum s_j}{1 + s_i \tanh \beta J \sum s_j}, \tag{7.49}$$

where the sum is over the nearest neighbors of s_i, and in the last step we used $\tanh(s_i x) = s_i \tanh x$ for $s_i = \pm 1$. Thus up to an overall constant that may be set to one, the transition rate for spin i is

$$w(s_i) = \frac{1}{2}\left[1 - s_i \tanh\left(\beta J \sum_j s_j\right)\right]. \tag{7.50}$$

At zero temperature, this rule forbids energy raising updates, while energy lowering updates occur with rate 1 and energy conserving events occur with rate 1/2. This defines a *majority rule* update — a spin flips to agree with the majority of its neighbors (Fig. 7.5). In contrast, the voter model is governed by *proportional rule* — a voter changes to the state of its local majority with a probability equal to the fraction of neighbors in this majority state. This proportionality allows one to factorize the voter model master equation in d dimensions into a product of soluble one-dimensional master equations. There is no such simplification for the Ising-Glauber model because s_j appears inside the hyperbolic tangent and the master equation is non-linear. For these reasons, much of our understanding of the Ising-Glauber model in $d > 1$ is based on simulation results or on continuum theories.

Another important feature of proportional rule is that there is no surface tension between domains of opposite-opinion voters. For example, a straight boundary between two opposite-opinion domains becomes fuzzier in voter model evolution (first line of Fig. 7.6). Additionally, even though the voter model coarsens in two dimensions, the lack of a surface tension means that the interface density disappears very slowly with time, namely, as $1/\ln t$. In contrast, for the Ising-Glauber model at zero temperature, there is a surface tension that scales as the inverse curvature for a droplet of one phase that is immersed in a sea of the opposite phase. We will discuss this surface tension in the next chapter; however, let us accept the existence of a surface tension that scales as the inverse curvature. Consequently, a single-phase droplet of radius R

[3] In $d = 1$, the two models are identical because the three distinct types of transitions of energy lowering, energy neutral, and energy raising,

$$\downarrow\underline{\uparrow}\downarrow \to \downarrow\underline{\downarrow}\downarrow \qquad \uparrow\underline{\uparrow}\downarrow \to \uparrow\underline{\downarrow}\downarrow \qquad \uparrow\underline{\uparrow}\uparrow \to \uparrow\underline{\downarrow}\uparrow$$

respectively, occur with the same rates of 1, 1/2, and 0.

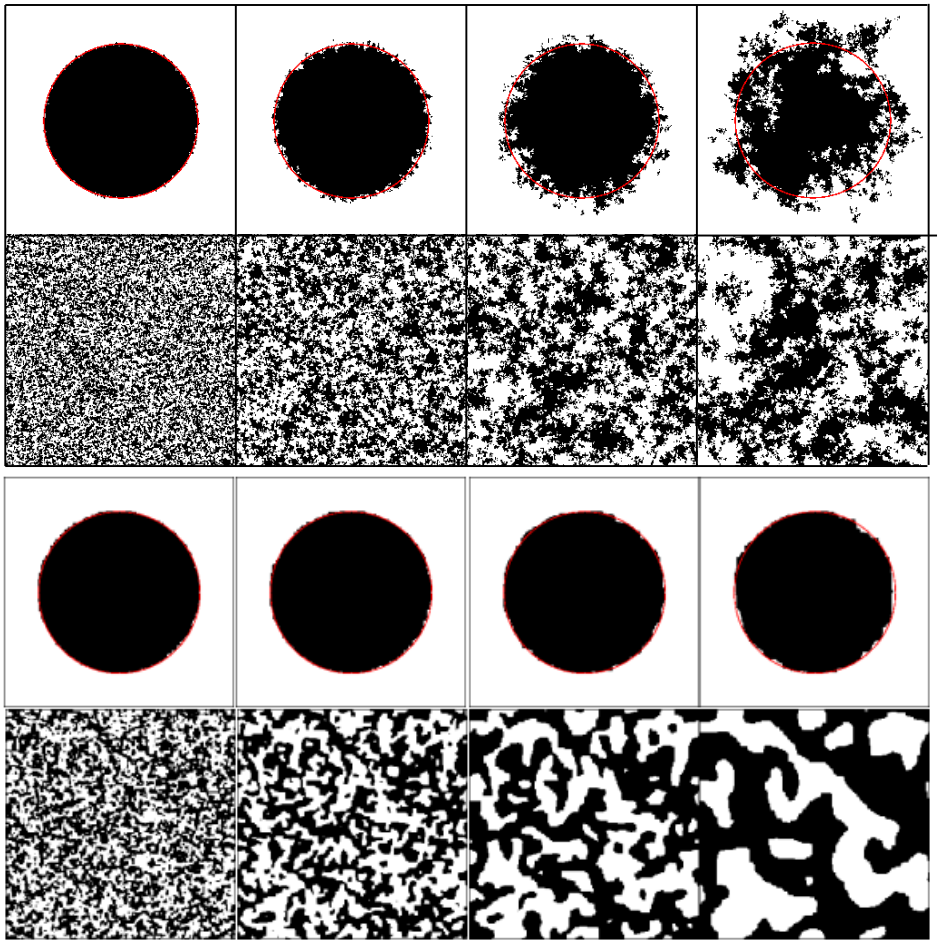


Figure 7.6: Spatial evolution in the voter model (top 2 rows) and the Ising-Glauber model at $T = 0$ (bottom two rows) on a 256×256 square lattice. Lines 1 & 3 shown snapshots at times $t = 4, 16, 64$, and 256 starting with an initial bubble of radius 180 for the voter model and the Ising-Glauber models, respectively. Lines 2 & 4 show the same evolution starting with a random initial condition with equal density of the two species. The voter model figure is from Dornic et al., Phys. Rev. Lett. **87**, 045701 (2001); courtesy of I. Dornic. The Ising-Glauber model figure is courtesy of V. Spirin.

in a background of the opposite phase will shrink according to $\dot{R} \propto -1/R$, or $R(t)^2 = R(0)^2 - at$ and thus disappear in a finite time (third line of Fig. 7.6). Additionally, the surface tension will quickly eliminate high curvature regions so that the coarsening pattern strongly differs from that of the voter model.

Perhaps the most basic questions about the Ising-Glauber model in $d > 1$ are concerned with the analog of the domain-size distribution. What is the nature of the coarsening when a system is prepared in a random initial state and then suddenly quenched to a low temperature? What is the final state? How long does it take to reach the final state? For $d > 1$ and for temperatures below the critical temperature, it has been well established that the system organizes into a coarsening domain mosaic of up and down spins, with the characteristic length scale that grows as $t^{1/2}$. For a finite system, this coarsening should stop when the typical domain size reaches the linear dimension L of the system.

However, when the final temperature T of the quench is strictly zero, intriguing anomalies occur when the size of the system is finite. At early stages of the relaxation, there is little difference in the dynamics of $T = 0$ and $T > 0$ systems. However, when the elapsed time is such that the characteristic time of the coarsening is comparable to the time to diffuse across the system, the two dynamics diverge. Perhaps the most striking feature of the $T = 0$ dynamics is that a system can get stuck in an infinitely long-lived metastable

state. These metastable states consist of straight stripes in two dimensions, but are more topologically more complex in higher dimension. In two dimensions, the probability of getting stuck in a metastable state is approximately $\frac{1}{3}$ as $L \to \infty$. In greater than two dimensions, the probability to reach the ground state rapidly vanishes as the system size increases. One obvious reason why the system fails to find the ground state is the rapid increase in the number of metastable states with spatial dimension. This proliferation of metastable states makes it more likely that a typical configuration will eventually reach one of these states rather than the ground state.

Mean field theory

Mean field theory describes systems in which fluctuations are negligible. One way to construct a mean-field theory for a spin system is to replace the actual environment surrounding each spin by the average environment that is then determined self consistently; this is the Curie-Weiss effective-field theory. Another mean-field description is achieved by embedding the Ising model on a *complete graph* of N sites, where all the $N(N-1)/2$ pairs of spin interact with the same strength. The Hamiltonian of the system is now

$$\mathcal{H} = -\frac{J}{N}\sum_{i<j} s_i s_j, \tag{7.51}$$

where the interaction strength scales inversely with the system size so that the energy is extensive, *i.e.*, scales linearly with N.

From the approach that gave the transition rate on a lattice in greater than one dimension [Eq. (7.50)], the transition rate for Glauber dynamics on the complete graph is simply

$$w(s_i) = \frac{1}{2}\left[1 - s_i \tanh\left(\frac{\beta J}{N}\sum_j s_j\right)\right]. \tag{7.52}$$

where the sum $\sum s_j$ is over all other spins in the system, from which the equation of motion for the mean spin is again $\frac{dS_i}{dt} = -2\langle s_i w_i\rangle$. We now exploit the fact that there are no fluctuations in the magnetization to write $\langle f(m)\rangle = f(\langle m\rangle)$. With this identity we have $\langle \tanh\frac{\beta}{N}\sum_i s_i\rangle = \tanh\frac{\beta}{N}\sum_i\langle s_i\rangle = \tanh\beta m$, with $m = N^{-1}\sum_i\langle s_i\rangle$ the average magnetization. Thus the equation of motion for the mean spin is

$$\frac{dS_i}{dt} = -S_i + \tanh\beta m. \tag{7.53}$$

Summing these rate equations, the average magnetization satisfies

$$\frac{dm}{dt} = -m + \tanh\beta m. \tag{7.54}$$

In contrast to one dimension, the magnetization is generally not conserved. The rate equation has three fixed points, one at $m = 0$ and two at $\pm m_{\rm eq}$, with the latter determined by the roots of the familiar transcendental equation $m = \tanh(\beta J m)$. A linear stability analysis shows that the zero-magnetization state is stable for $\beta J \le 1$ but unstable for $\beta J > 1$, and vice versa for the states with $m = \pm m_{\rm eq}$. Thus there is a phase transition at $\beta_c J = 1$, with $m_{\rm eq} \sim (T_c - T)^{1/2}$ for $T \lesssim T_c$. The emergence of two equivalent, but symmetry-breaking ground states when the Hamiltonian itself is symmetric is termed *spontaneous symmetry breaking*.

Above the critical temperature, the magnetization decays to zero and we expand $\tanh\beta m$ in Eq. (7.54) in powers of βm to give

$$\frac{dm}{dt} = -(\beta_c - \beta)m - \frac{1}{3}(\beta m)^3. \tag{7.55}$$

In the high temperature phase, the cubic term is negligible so that the magnetization decays exponentially in time, $m \sim e^{-t/\tau}$, with $\tau = (\beta_c - \beta)^{-1}$. At the critical point, the relaxation is algebraic,

$$m \sim t^{-1/2}. \tag{7.56}$$

Below the critical temperature, the magnetization also decays exponentially toward its equilibrium value, $|m-m_{\rm eq}| \sim e^{-t/\tau}$, with $\tau^{-1} = 1-\beta/\cosh^2(\beta m_{\rm eq})$ **not explained**. Thus, as the critical point is approached, either from above or from below, the relaxation time scale diverges as

$$\tau \sim |T_c - T|^{-1}. \tag{7.57}$$

The divergence of the relaxation time as $T \to T_c$ is a generic sign of critical slowing down as the approach to equilibrium becomes extremely slow.

7.4 Kawasaki Spin-Exchange Dynamics

The transition rate

As mentioned at the outset of Sec. 7.3, there are two fundamental classes of spin dynamics: magnetization conserving and magnetization non-conserving. The former class is appropriate to describe alloy systems, where the two different spin states naturally correspond to the two component atoms that comprise the alloy. In studying the dynamics of phase separation of an alloy into domains of pure metal, a plausible dynamics is that the positions of atoms of different species are exchanged; there is no alchemy where one type of atom can be converted to the other type. In this section, we investigate a simple realization of this order-parameter conserving dynamics — Kawasaki dynamics.

In Kawasaki dynamics, neighboring antiparallel spins simultaneously reverse their states so that

$$\cdots \underline{\uparrow\downarrow} \cdots \longrightarrow \cdots \underline{\downarrow\uparrow} \cdots . \tag{7.58}$$

Alternatively, the two spins can be regarded as being exchanged and hence the term *spin-exchange*. Clearly, such moves do not alter the magnetization and so this quantity is *strictly* conserved in every update event. The existence of this strict conservation law has far-reaching consequences that will become clearer when we discuss continuum theories of spin dynamics in the next chapter.

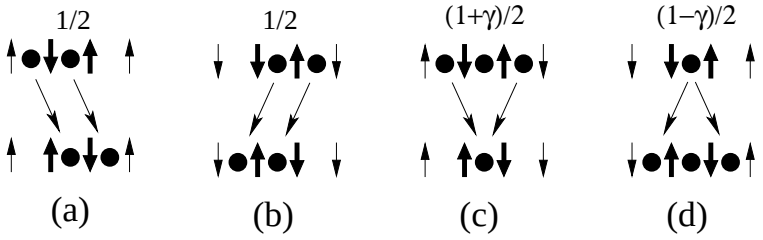


Figure 7.7: Energy neutral update events (a) & (b), energy lowering events (c), and energy raising events (d) for Kawasaki dynamics in one dimension. The spins that flip are shown bold. Also shown are the corresponding domain wall particles (•) and the transition rates for these four events.

Again, there are three types of update events: energy raising, energy lowering, and energy neutral. As illustrated in Fig. 7.7, the energy neutral update is equivalent to the simultaneous hopping of two nearest-neighbor domain walls, or to the hopping of an impurity down spin in a sea of up spins. As long as the domain-wall pair remains isolated from all other domain walls, the pair hops freely between neighboring lattice sites. This pair can be viewed as an elementary excitation of the spin system. The hopping rate of a domain wall pair merely sets the time scale, so there is no loss of generality in setting this rate to 1/2, as in Glauber dynamics. Because such diffusive moves do not alter the energy, they automatically satisfy the detailed balance condition. The rates of the remaining two update events are then set by detailed balance. Since spin exchange involves the interactions among four spins—the two spins that flip and their two neighbors—the rates depend on the total energy of the three bonds connecting these four spins. The detailed balance condition is then

$$\frac{w_3}{w_{-1}} = \frac{p_{-1}}{p_3} = \exp(4\beta J), \tag{7.59}$$

where w_q is the transition rate out of a state with energy qJ and p_q its equilibrium probability. Using the convenient Glauber notations of $w_3 = (1+\gamma)/2$ and $w_{-1} = (1-\gamma)/2$ for energy raising and energy lowering transitions, the detailed balance condition has the same form as in Glauber dynamics, $\frac{1+\gamma}{1-\gamma} = \exp(4\beta J)$, or $\gamma = \tanh 2\beta J$.

To determine the transition rates, we first must guarantee that spins i and $i+1$ are antiparallel. This constraint can be achieved by the factor $(1 - s_i s_{i+1})/2$ that equals $+1$ if the two spins are antiparallel and equals zero otherwise. The form of the flip rate then depends on the interaction energy between the pairs s_{i-1} and s_i, and between s_{i+1} and s_{i+2}. The flip rate should be a symmetric function of these two bond energies and the rate should be proportional to $(1+\gamma)/2$, $1/2$, and $(1-\gamma)/2$ respectively, when the signs of these bond energies are $--$, $+-$, and $++$. These constraints leads to the transition rate

$$w_i(s_i, s_{i+1}) = \frac{1}{2}\left[1 - \frac{\gamma}{2}(s_{i-1}s_i + s_{i+1}s_{i+2})\right] \times \frac{1}{2}(1 - s_i s_{i+1}). \tag{7.60}$$

An important feature of this rate is that the evolution of spin correlation functions are no longer closed. One-spin averages are coupled to three-spin averages, two-spin averages are coupled to four-spin averages, *etc.* Thus the equation of motion for a particular correlation function generates an infinite hierarchy of equations for high-order correlations. This coupling to higher-order correlation functions arises in a wide range of many-body problems and it is a matter of considerable technical effort and artistry to find a tractable and accurate scheme to truncate this infinite hierarchy.

Frustration at zero temperature

Because Kawasaki dynamics is more constrained than Glauber dynamics, a system will almost always get stuck forever at zero temperature in one of the very large number of metastable states — one whose energy is above the ground state and for which the only possible transitions by Kawasaki dynamics would raise the energy (see Fig. 7.7(d)). The metastable states are characterized by each domain wall particle being separated by more than a nearest-neighbor distance from any other domain wall. Equivalently the lengths of all spin domains are two or longer. The number of such configurations in a system of length L asymptotically grows as g^L, where $g = (1+\sqrt{5})/2$ is the golden ratio. It is striking how often this beautiful number appears in statistical physics problems. At zero temperature these metastable states prevent the system from reaching the ground state. At non-zero temperature, these states merely slow the approach toward equilibrium.

To study how the system evolves to a metastable state, let's study the case where energy lowering transitions only are allowed, as illustrated in Fig. 7.7(c). The resulting behavior differs only slightly from the situation where diffusive moves are also allowed, but the former is simpler to treat analytically. The dynamics is perhaps best visualized in terms of the domain walls that occupy the sites of the dual lattice. According to Fig. 7.7(c), an update step consists of picking three contiguous domain wall particles at random and then removing the two outside particles. Since pairs of domain walls are removed sequentially from triplets of consecutive domain walls, the process is equivalent to the random sequential adsorption of a $\bullet\circ\bullet$ "fork" on top of a string of three consecutive domain wall particles. Because of this equivalence, we can use the tools of random sequential adsorption (Chapter 6) to solve the problem.

Let E_k be the probability that a string of k sites (in the dual lattice) are all occupied by domain walls. This probability evolves by the master equation

$$\frac{dE_k}{dt} = -(k-2)E_k - 2E_{k+1} - 2E_{k+2} \tag{7.61}$$

for $k \geq 3$. This equation reflects the different ways that the transition $\circ\circ\circ \to \bullet\circ\bullet$ can occur and alter the number of empty strings of length k. There are $k-2$ ways that this transition can occur in the interior of a k-string. There are also 2 ways that this transition can occur with two sites at the edge of the k-string and one site outside, and also 2 ways with one site at the edge of the k-string and two sites outside.

We solve this rate equation by introducing the exponential ansatz $E_k = \Phi(t)\, e^{-(k-2)t}$ [see the discussion accompanying Eq. (6.4)]. For the initial condition of an antiferromagnetic spin state, the dual lattice is completely occupied. Thus $E_k = 1$ initially, so that $\Phi(0) = 1$. Substituting this ansatz into the rate equation (7.61) leads to the ordinary differential equation for Φ:

$$\frac{d\phi}{dt} = -2\phi(e^{-t} + e^{-2t}). \tag{7.62}$$

Integrating this equation gives the string probabilities for $k \geq 2$,

$$E_k(t) = \exp\left[-(k-2)t + e^{-2t} + 2e^{-t} - 3\right]. \tag{7.63}$$

Since two domain walls are lost in each update event and these events occur with rate E_3, the domain wall density $\rho \equiv E_1$ satisfies $\frac{d\rho}{dt} = -2E_3$. Using Eq. (7.63) for E_3 and integrating then yields the domain wall density

$$\rho(t) = 1 - 2\int_0^t ds \exp\left[-s + e^{-2s} + 2e^{-s} - 3\right]. \tag{7.64}$$

The final "jamming" density is finite, $\rho_{\rm jam} \equiv \rho(\infty) = 0.450898\ldots$. Thus there is not very much relaxation as almost half of the domain walls still remain in the final jammed state. Moreover, the relaxation to the jamming density is exponential in time,

$$\rho(t) - \rho_{\rm jam} \simeq e^{-3}e^{-t}. \tag{7.65}$$

We see that the system neither reaches the lowest energy state, nor does it exhibit critical slowing down. The underlying reason for both of these behaviors is that the dynamics samples only a very restricted portion of the phase space.

Coarsening at infinitesimal temperature

While the one-dimensional chain with Kawasaki dynamics quickly reaches a jammed state at zero temperature, the equilibrium state will be reached for any non-zero temperature, no matter how small. Because the correlation length diverges as the temperature approaches zero, one can set the temperature sufficiently small so that the correlation length is much larger than the length of the system. Then the equilibrium state consists of a single domain and we are interested in the approach to this final state.

The large separation of time scales between energy raising updates and all other update events leads to an appealing description of the domain evolution within the framework of an extremal dynamics. Since the rate of an energy raising update equals $e^{-4\beta J}$, the typical time for such an event is $\tau \equiv e^{4\beta J}$. We define the time unit to be $e^{4\beta J}$. Energy neutral and energy lowering events then occur instantaneously in this time unit. Starting from an initial state, the system instantly reaches a frustrated state in which no further energy neutral or energy lowering moves are possible. After a time τ has elapsed (on average) an energy raising event occurs that is then followed by a burst of energy neutral and energy lowering events until the system reaches another frustrated state. This pattern of an energy raising event followed by a burst of complementary events continues until a finite system reaches the ground state. As we will show, this dynamics leads to the typical domain size growing in time as $t^{1/3}$ and is a general feature of order-parameter conserving dynamics. One of the appealing features of Kawasaki dynamics in one dimension is that this $t^{1/3}$ coarsening emerges in a direct way. In contrast, we will see in the next chapter that it is much more subtle to deduce the $t^{1/3}$ coarsening from continuum approaches.

At long times, the system evolves to a low-energy state that consists of alternating domains of typical length ℓ. The subsequent evolution at low temperature is controlled by rare, energy raising updates where a pair of domain walls nucleates around an existing isolated domain wall. Once this triplet of domain walls forms, a bound pair of these domain walls can diffuse freely with no energy cost until another isolated domain wall is encountered. When such a collision occurs, two of the domain walls annihilate so that a static single domain wall remains. As illustrated in Fig. 7.8, the creation of a mobile bound domain wall pair is equivalent to an isolated spin splitting off from a domain and then diffusing freely within a neighboring domain of length ℓ of the opposite orientation. If this diffusing spin returns to its starting point, the net effect is no change in the domain configuration. However, if the spin manages to traverse to the other side of the domain, then one domain has increased its size by one and another has shrunk by one. This effective diffusion of domain lengths is the mechanism that drives the coarsening.

What is the probability that the spin can actually traverse to the other side of the domain? This is given by the classic "gambler's ruin" problem as discussed in the highlight on page 23. Once the spin has split off, it is a distance 1 from its initial domain and a distance $\ell - 1$ from the domain on the other side. Since the spin diffuses freely, it eventually reaches the other side with probability $1/\ell$, while the spin returns to its

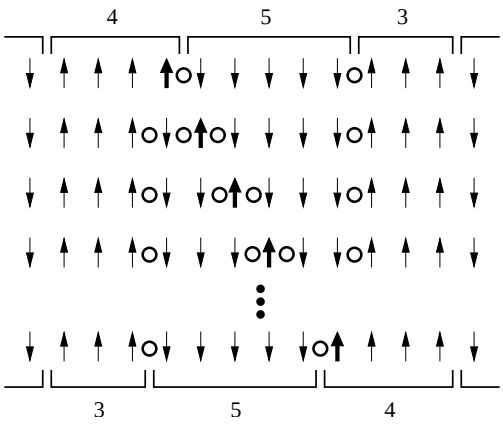


Figure 7.8: Illustration of the effective domain diffusion from Kawasaki dynamics at infinitesimal temperature. The second line shows an energy raising event where a spin (shown bold) splits off from a domain. Eventually this spin joins the next domain to the right. Also shown is the evolution of the domain walls. The net result of the diffusion of the spin across the middle domain is that this moves one step to the left.

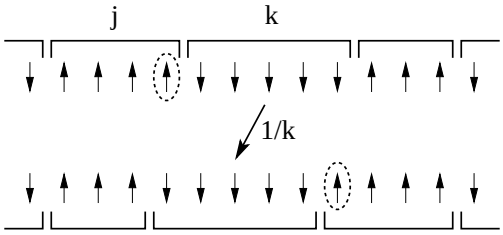


Figure 7.9: Effective domain diffusion by Kawasaki dynamics at infinitesimal temperature. A $\downarrow$ spin from the j domain (dashed oval) splits off and eventually reaches the right edge of the k domain. The k domain has moved rigidly to the left by one lattice spacing.

starting position with probability $1 - 1/\ell$. Thus the probability that the ℓ-domain hops by one step equals $1/\ell$. That is, the diffusion coefficient of a domain equals the inverse of its length: $D(\ell) = \ell^{-1}$.

Thus in the low-temperature limit, the spin dynamics maps to an effective isotropic hopping of entire domains by one step to the left or the right[4] (Fig. 7.9). Domains of length 1 disappear whenever one of their neighboring domain hops toward them. Concomitantly, the lengths of the neighboring domains are rearranged so that four domains merge into two (Fig. 7.10). The net effect of these processes is coarsening because domains of length 1 disappear. We can determine the typical domain length by a heuristic argument. Because each domain performs a random walk, coalescence occurs whenever a domain diffuses of the order of its own length. In such a coalescence, a domain typically grows by an amount $\Delta\ell$ that is also of the order of ℓ, while the time between coalescences is $\Delta t \sim \ell^2/D(\ell)$. Thus

$$\frac{\Delta\ell}{\Delta t} \sim \frac{\ell}{\ell^2/D(\ell)} \sim \frac{1}{\ell^2} \ ,$$

so that domains grow as

$$\ell \sim t^{1/3}. \tag{7.66}$$

It is conventional to define the *dynamical exponent* z in terms of the growth of the typical length scale in a coarsening process via $\ell \sim t^z$. For the non-conserved Glauber and the conserved Kawasaki dynamics,

[4]There is an anomaly involving domains of length 2 that can be ignored for the purposes of this discussion.

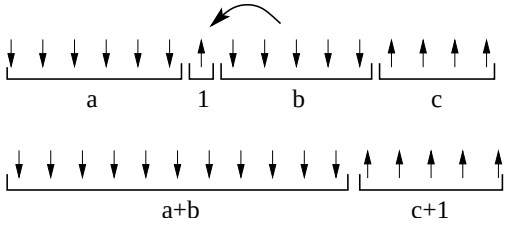


Figure 7.10: The outcome after domain merging.

the dynamical exponent is:

$$z = \begin{cases} 1/2 & \text{nonconservative dynamics,} \\ 1/3 & \text{conservative dynamics.} \end{cases} \tag{7.67}$$

While we have derived these results in one dimension, they are generic for all spatial dimensions. Conservation laws are a crucially important ingredient in determining the nature of non-equilibrium dynamics.

7.5 Cluster Dynamics

Glauber single-spin flip dynamics and the Kawasaki spin-exchange dynamics are local in that they involve flipping a single spin or a pair of spins. In spite of their idealized natures, these rules were the basis of many simulational studies of coarsening and dynamic critical phenomena because of their connection to the evolution of real systems. However, a dynamics that is based on flipping single spins is computationally inefficient. Compounding this inefficiency, the dynamics significantly slows down close to criticality. To mitigate these drawbacks, Swendsen and Wang developed a dynamical update rule in which an entire suitably-defined cluster of spins is flipped simultaneously. Because of their efficiency, cluster algorithms have been used extensively to simulate the equilibrium behavior of many-body statistical mechanical and lattice field theory models. The Swendsen-Wang and the Wolff algorithms are two of the earliest and most prominent such examples of cluster dynamics. Remarkably, both of these algorithms are analytically soluble by the master equation approach.

Swendsen-Wang dynamics

In one dimension, an Ising spin chain consists of alternating spin-up and spin-down domains. In the Swendsen-Wang algorithm, an entire domain of aligned spins is chosen at random and all its spins are flipped simultaneously, as illustrated below:

$$\cdots\uparrow\underbrace{\downarrow\downarrow\downarrow\downarrow\downarrow}\uparrow\cdots\longrightarrow\cdots\uparrow\underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow}\uparrow\cdots.$$

By construction, all such updates decrease the energy. In each update event, there is a net loss of two domains. Consequently, the number density of domains ρ decreases according to $\frac{d\rho}{dt} = -2\rho$, where we take the flip rate to be 1, without loss of generality. The density of domains then decreases exponentially with time, and for the antiferromagnetic initial condition in which $\rho(0) = 1$, the domain density is $\rho(t) = e^{-2t}$. Since the average domain length $\langle k\rangle$ is the inverse of the domain density, $\langle k\rangle = e^{2t}$. When this average length reaches the system length L the dynamics is complete. This criterion yields the time to reach the ground state $T_L \propto \ln L$.

Now consider the domain length distribution. We define c_ℓ as the density of domains of length ℓ. When a domain is flipped, it merges with its two neighbors, so that the length of the resulting domain equals the length of these three domains. As a result of this three-body aggregation process, c_ℓ evolves according to

$$\frac{dc_\ell}{dt} = -3c_\ell + \frac{1}{\rho^2}\sum_{i+j+k=\ell} c_i\, c_j\, c_k. \tag{7.68}$$

The factor of $-3c_\ell$ accounts for the loss of a domain that occurs when a domain of length ℓ or either of its neighboring domains is flipped. The last term accounts for the gain in c_ℓ due to the flipping of a domain of length j that then merges with its two neighboring domains of lengths i and k, with $\ell = i + j + k$. The simplest way to deduce the prefactor ρ^{-2} is to ensure this master equation consistent with the rate equation for the domain density $\dot\rho = -2\rho$. Notice that newly-created domains do not affect their neighbors, nor are they affected by their neighbors. Thus if the domains are initially uncorrelated, they remain uncorrelated. Because no spatial correlations are generated, the rate equations are exact!

We can obtain a cleaner-looking master equation by introducing $P_\ell \equiv c_\ell/\rho$, namely, the probability for a domain of length ℓ (with the normalization $\sum_\ell P_\ell = 1$). Using $\dot\rho = -2\rho$ in Eq (7.68), P_ℓ evolves as

$$\frac{dP_\ell}{dt} = -P_\ell + \sum_{i+j+k=\ell} P_i P_j P_k. \tag{7.69}$$

As we have seen in many previous examples, the convolution form of the gain term cries out for applying the generating function method. Thus we introduce the generating function $F(z) = \sum_\ell P_\ell\, z^\ell$ into (7.69) and find that it satisfies $\frac{\partial F}{\partial t} = -F + F^3$. Writing $1/(F^3 - F)$ in a partial fraction expansion,t the equation can be integrated by elementary methods and the solution is

$$F(z,t) = \frac{F_0(z)e^{-t}}{\sqrt{1 - F_0(z)^2(1 - e^{-2t})}}, \tag{7.70}$$

where $F_0(z)$ is the initial generating function.

For the antiferromagnetic initial condition, the initial condition is $F_0(z) = z$. Expanding the generating function in powers of z then yields the domain number distribution

$$P_{2\ell+1} = \binom{2\ell}{\ell}\left(\frac{1 - e^{-2t}}{4}\right)^\ell e^{-t} \tag{7.71}$$

in which domains have odd lengths only. Since the average domain length grows exponentially with time, $\langle\ell\rangle = e^{2t}$, we expect that this scale characterizes the entire length distribution. Employing Stirling's approximation, we find that asymptotically the length distribution approaches the scaling form $P_\ell \to e^{-2t}\Phi(\ell e^{-2t})$ with the scaling function

$$\Phi(x) = \frac{1}{\sqrt{2\pi x}}\, e^{-x/2}. \tag{7.72}$$

Because the scaling function diverges $\Phi(x) \sim x^{-1/2}$ for $x \ll 1$, there is a large number of very small domains.

Wolff dynamics

In the Wolff cluster algorithm, a spin is selected at random and the domain it belongs to is flipped. This protocol further accelerates the dynamics compared to the Swendsen-Wang algorithm because the larger the domain, the more likely it is updated. Schematically, the Wolff dynamics is

$$\cdots\uparrow\underbrace{\downarrow\downarrow\cdots\downarrow\downarrow}_{k}\uparrow\cdots \xrightarrow{k} \cdots\uparrow\underbrace{\uparrow\uparrow\cdots\uparrow\uparrow}_{k}\uparrow\cdots, \tag{7.73}$$

so that a flipped domain again simply merges with its neighbors. Since each spin is selected randomly, the time increment associated with any update is identical. The domain density therefore decreases with constant rate $\dot\rho = -2$, so that $\rho(t) = 1 - 2t$ and the entire system is transformed into a single domain in a finite time, $t_c = 1/2$. Correspondingly, the average domain length, $\langle k\rangle = (1 - 2t)^{-1}$, diverges as $t \to t_c$.

The evolution of the domain length distribution is governed by the natural generalization of (7.69)

$$\frac{dP_\ell}{dt} = -\ell P_\ell + \sum_{i+j+k=\ell} j P_i P_j P_k. \tag{7.74}$$

The generating function $F(z,t) = \sum_\ell P_\ell z^\ell$ satisfies

$$\frac{\partial F}{\partial t} = z\,(F^2 - 1)\,\frac{\partial F}{\partial z}. \tag{7.75}$$

To solve this equation, we first transform from the variables (t,z) to $(\tau,y) \equiv (t, t - \ln z)$ to absorb the negative term on the right-hand side. This transformation gives

$$\frac{\partial F}{\partial \tau} = -F^2\,\frac{\partial F}{\partial y}. \tag{7.76}$$

We now employ the same procedure as that used in the solution of aggregation with the product kernel (see the discussion leading up to Eq. (4.37) in chapter 4) to transform among the variables (τ, y, F) and reduce (7.76) into the linear differential equation $\frac{\partial y}{\partial \tau} = F^2$. The solution to this equation is $y = G(F) + F^2\tau$, with $G(F)$ determined by the initial conditions, or, equivalently,

$$t - \ln z = G(F) + F^2 t. \tag{7.77}$$

For the antiferromagnetic initial condition $F_0(z) = z$, so that $G(F) = -\ln F$. Substituting $G(F) = -\ln F$ into (7.77) and exponentiating yields the following implicit equation for the generating function

$$z = F\,e^{t - F^2 t}\,. \tag{7.78}$$

The length distribution P_k is just the k^{th} term in the power series expansion of $F(z)$. Formally, this term may be extracted by writing P_k in terms of the contour integral

$$P_k = \frac{1}{2\pi i}\oint \frac{F(z)}{z^{k+1}}\,dz\,,$$

then transforming the integration variable from z to F, and using the Lagrange inversion formula (see the discussion on page 51 in Chapter 4). These steps give

$$\begin{aligned} P_k = \frac{1}{2\pi i}\oint \frac{F(z)}{z^{k+1}}\,dz &= \frac{1}{2\pi i}\oint \frac{F}{z(F)^{k+1}}\,\frac{dz}{dF}\,dF, \\ &= \frac{e^{-kt}}{2\pi i}\oint e^{kF^2 t}\left[\frac{1}{F^k} - \frac{2t}{F^{k-2}}\right]\,dF, \end{aligned} \tag{7.79}$$

where we use the fact that $\frac{dz}{dF} = e^{t-F^2 t}(1 - 2F^2 t)$ in the above integral. Now we find the residues simply by expanding $e^{kF^2 t}$ in a power series and keeping only the coefficient of $\frac{1}{F}$ in the integrand. Because the power series is even in F, only P_k for odd values of k is non zero and we find

$$P_k = e^{-kt}\left[\frac{(kt)^{(k-1)/2}}{\left(\frac{k-1}{2}\right)!} - 2t\frac{(kt)^{(k-3)/2}}{\left(\frac{k-3}{2}\right)!}\right].$$

After some simple algebra, the domain length distribution is

$$P_{2k+1}(t) = \frac{(2k+1)^{k-1}}{k!}\,t^k\;e^{-(2k+1)t}\,. \tag{7.80}$$

Near the critical time ($t \to 1/2$), Stirling's approximation gives, for the leading behavior of domain length distribution,

$$P_k(t) \simeq \frac{1}{\pi^{1/2}\,k^{3/2}}\;e^{-k\,\rho^2/4}\;, \tag{7.81}$$

with $\rho = 1 - 2t$. While this distribution has a characteristic length scale $k^* = 4/\rho^2 = (1/2 - t)^{-2}$, this length does not fully characterize the distribution. Since the domain length distribution has a power-law $k^{-3/2}$ tail that is cut off at $k^* \propto \rho^{-2}$, the average domain length $\langle k\rangle = \sum_k k\,P_k \sim \int^{k^*} k\,k^{-3/2}\,dk$. This last integral then gives the expected behavior $\langle k\rangle \sim \rho^{-1}$.

Problems

7.1 The Voter Model

1. Evaluate the average opinion for a Democrat in a sea of uncommitted voters: $S(\mathbf{x}, 0) = \delta(\mathbf{x})$.

7.2 & 7.3 Glauber Spin-Flip Dynamics

1. Solve for the domain number distribution in the Potts model with Glauber dynamics. Hint: P_1 is replaced by $P_1/(q-1)$ in Eq. (7.43).
2. Verify that for the Hamiltonian $\mathcal{H} = -\sum_{i<j} J_{i,j} s_i s_j$ the spin flip rate is $w_i = \frac{1}{2}\big(1 - \tanh \beta s_i \sum_j J_{i,j} s_j\big)$.
3. Obtain the entropy in the mean-field model.
4. Determine $P_M(t)$, the probability to have M up spins and $N-M$ down spins,[5] for zero-temperature dynamics.
5. Examine $P_M(t)$ for the critical dynamics.
6. In the low temperature regime ($\infty > \beta > \beta_c$), the distribution $P_M(t)$ is bimodal with peaks of width $\propto \sqrt{N}$ around $M_\pm = \frac{1}{2} N (1 \pm m_\infty)$. The system spends almost all time in the proximity of the peaks yet occasionally it leaves one peak and reaches the other. Estimate the transition time.

7.4 Glauber Spin-Exchange Dynamics

1. Show that the correlation functions obey an infinite hierarchy of equations. Write the evolution equation for S_i.
2. Obtain the number of frozen configurations in the zero-temperature limit for Kawasaki dynamics.
3. Solve for the domain wall density at zero-temperature for random initial conditions.

7.5 Cluster Dynamics

1. Consider the zero-temperature Swendsen-Wang dynamics, with energy lowering moves only, for the $q = \infty$ Potts model. In this case a domain merges with only one of its neighbors. Determine the domain density and the domain length distribution.

   Solution: Since there is a net loss of one domain in a single update, the number density obeys $d\rho/dt = -\rho$. Therefore $\rho(t) = e^{-t}$ while the average domain length again grows exponentially with time. The domain length size distribution now evolves by two-body aggregation, so that this distribution satisfies

   $$\frac{dP_k}{dt} = -P_k + \sum_{i+j=k} P_i P_j. \tag{7.82}$$

   To solve this equation, we again introduce the generating function into this equation to give $\frac{\partial F}{\partial t} = -F + F^2$, whose solution is simply

   $$F(z,t) = \frac{F_0(z) e^{-t}}{1 - (1 - e^{-t}) F_0(z)}. \tag{7.83}$$

   Expanding this generating function in a power series we immediately obtain

   $$P_k(t) = e^{-t}(1 - e^{-t})^{k-1}. \tag{7.84}$$

   Asymptotically, the distribution attains the scaling form $P_k(t) \sim e^{-t}\Phi(ke^{-t})$ with the purely exponential scaling function $\Phi(x) = \exp(-x)$. The enhancement of smaller than average domains disappears in the $q \to \infty$ limit.

2. Analyze a domain coarsening process where the smallest domain merges with one of its neighbors.

[5] In all problems in this paragraph the system is *finite*.

Chapter 8

COARSENING

In the previous chapter, we discussed the coarsening of kinetic Ising models starting from an initial homogeneous high-temperature phase and suddenly lowering the temperature (quenching) below the critical temperature. The ensuing coarsening is complex (see Fig. 7.6), and considerable effort has been devoted to developing simpler continuum descriptions. While we lose a direct connection to individual spins in such a formulation, continuum theories, while still quite formidable, are typically simpler than their discrete counterparts. We will investigate some of the simplest geometries for coarsening dynamics that can be solved explicitly. This includes freely diffusing interfaces in the absence of interactions, the single domain wall, and a droplet of one phase that is immersed into another phase. For such cases, many aspects of the evolution can be found analytically. For the more generic case of a random initial condition, the continuum approach provides many new insights that seem to be impossible to obtain by a description at the level of individual spins.

8.1 The Models

We tacitly assume that the order parameter is a scalar, and will state otherwise explicitly. We generally have in mind magnetic systems and will use the terminology associated with such systems. However, this usage is more reflective of tradition rather than the dominant application of coarsening. Even for the simplest case of systems with a scalar order parameter, the distinction between non-conservative and conservative dynamics is crucially important, and we now outline the generic models of these two dynamics.

Non-conservative dynamics

The basic ingredients that underlie non-conservative Landau-Ginzburg dynamics are the following:

- We deal with a continuous coarse-grained order parameter, or magnetization, $m(\mathbf{x},t)$ rather than a binary Ising variable $\sigma = \pm 1$ by defining $m(\mathbf{x},t) \equiv \ell^{-d}\sum\sigma$. That is, $m(\mathbf{x},t)$ is the average magnetization in a block of linear dimension ℓ; the sum runs over the ℓ^d spins in the block centered at $\mathbf{x}$. Here ℓ should be much greater than the lattice spacing a and much smaller than the system size $\mathcal{L}$. These restrictions give a coarse-grained magnetization that varies smoothly on a scale greater than ℓ that can accommodate a non-trivial spatial dependence. This coarse-graining is appropriate over the time range where the typical domain size is much larger than ℓ and much smaller than $\mathcal{L}$.

- We describe the thermodynamics of the system by the coarse-grained Landau free-energy functional

$$F[m(\mathbf{x})] = \int \left[\frac{1}{2}\,|\nabla m(\mathbf{x})|^2 + V(m(\mathbf{x}))\right] d\mathbf{x}, \tag{8.1}$$

  where the potential $V(m)$ has a double-well structure with minima corresponding to the equilibrium states. The standard example is $V(m) = \frac{1}{2}\,(1-m^2)^2$ for systems with two degenerate minima; we always assume that the minima occur at $m = \pm 1$ and that $V(\pm 1) = 0$. While this coarse-grained

free energy is intuitively plausible — it combines the Landau mean-field theory with the lowest order contribution due to spatial variation of the order parameter — it is not possible to derive Eq. (8.1) from first principles starting with a microscopic model, such as the Ising model.

- The final step in constructing a coarse-grained dynamical description is also phenomenological: we simply assert that the order parameter changes at a rate that is proportional to the local thermodynamic force $\frac{\delta F}{\delta m(\mathbf{x},t)}$, where $\frac{\delta}{\delta m(\mathbf{x},t)}$ denotes the functional derivative with respect to m at position $\mathbf{x}$. Thus we are considering *overdamped* dynamics because there is no second time derivative in the equation of motion. This assumption is equivalent to ignoring inertia. Absorbing a coefficient of proportionality into the time scale, we thus arrive at the *time-dependent Ginzburg-Landau (TDGL) equation*

$$\frac{\partial m}{\partial t} = -\frac{\delta F}{\delta m} = \nabla^2 m - V'(m). \tag{8.2}$$

The TDGL equation is one of the central equations of coarsening.

There are three important aspects of the TDGL equation that merit emphasis. First, even though the TDGL equation is widely accepted and used, this dynamics is impossible to derive from a microscopic theory, such as the Ising-Glauber model. Second the TDGL equation is purely dissipative. This feature can be easily shown by computing the change in the free energy as a function of time:

$$\frac{dF}{dt} = \int \frac{\delta F}{\delta m}\,\frac{\partial m}{\partial t}\,d\mathbf{x} = -\int \left(\frac{\delta F}{\delta m}\right)^2 \mu d\mathbf{x} \leq 0.$$

Thus a system governed by the TDGL equation simply "flows" down the free energy gradient until one of the two potential minima is reached. A third important point is that the TDGL equation has a natural interpretation as a reaction-diffusion process. For the standard potential $V(m) = \frac{1}{2}\,(1-m^2)^2$, the explicit TDGL equation is:

$$\frac{\partial m}{\partial t} = \nabla^2 m + 2m(1-m^2). \tag{8.3}$$

If we view m as a density, then the above equation describes the evolution of a diffusing population of particles that give birth, $m \to 2m$, and undergo 3-body coalescence, $3m \to m$. The rate equation for this reaction, $\frac{dm}{dt} = 2m - 2m^3$, has an unstable fixed point at $m = 0$ and a stable fixed point at $m = 1$. When diffusion is included, the resulting equation of motion (8.3) describes the infiltration of a stable high-density phase ($m = 1$) into a low-density ($m = 0$) region. In fact, Eq. (8.3) is just one example of the family of self-regulating reactions of the form

$$\frac{\partial m}{\partial t} = \nabla^2 m + f(m), \tag{8.4}$$

with $f(0) = f(1) = 0$, $f(m) > 0$ for $0 < m < 1$, and $f(m)$ having a single maximum in the range $[0, 1]$. The most famous of this class of equations arises when $f(m) = m(1-m)$, and is known as the Fisher-Kolmogorov-Petrovsky-Piscounov (FKPP) equation (see the highlight below).

The FKPP Equation

The FKPP equation

$$\frac{\partial A}{\partial t} = D\nabla^2 A + kA(N-A), \tag{8.5}$$

describes the evolution of a diffusing population where individuals give birth ($A \to A + A$) and quadratically self-regulate ($A + A \to A$). If spatial fluctuations are neglected, then in the resulting rate equation, $\dot{A} = A(N - A)$, there is a stable fixed point when $A = N$ and an unstable fixed point for $A = 0$. Here N should be viewed as the maximum number of particles that can be accommodated on a single site in the lattice version of the reaction. Starting with any non-zero density in a finite system, a final density of N is reached. An important example for population biology is the case where $A = N$ in the half-space $x < 0$ and $A = 0$ for $x > 0$. By diffusion alone, the sharp interface between these two regions would become smoother. However, because of the interplay between diffusion and reaction, the stable high-density phase propagates as a stationary wave into the low-density phase.

Let's study the basics of this wave propagation in one dimension. Proceeding in a naive way, we assume that the stable state infiltrates as a stationary propagating wave. As a preliminary, it is useful to rescale the density by $A \to A/N$, time by $t \to kt$, and then the length by $x \to x\sqrt{k/D}$ to non-dimensionalize the FKPP equation to $\frac{\partial A}{\partial t} = \nabla^2 A + A(1-A)$. The simplest assumption is that the leading edge of this wave decays exponentially with position:

$$A(x,t) \sim e^{-\lambda(x-vt)}\,.$$

Substituting this form into Eq. (8.5) and dropping the non-linear term that is negligible at the leading edge of the wave, we obtain the dispersion relation

$$v = \lambda + \frac{1}{\lambda}.$$

Thus the velocity ostensibly depends on λ and we can only surmise that $v \geq 2$. The equation of motion does not provide any information about which velocity is selected. Nature appears to be wise, however, and typically "selects" the minimum velocity $v_{\min} = 2$ that arises when $\lambda = 1$. In fact, for most "reasonable" initial conditions, this minimum velocity is the one that is actually selected.

More precisely, if the initial density decays as $e^{-\lambda_0 x}$ with $\lambda_0 > 1$, then the leading edge of the wave asymptotically evolves to an exponential with decay parameter $\lambda = 1$ and velocity $v = v_{\min}$. We call such wavefronts "sharp". On the other hand, if $\lambda_0 < 1$, then the leading edge of the wave preserves this slower decay, $e^{-\lambda_0(x-vt)}$, and the velocity is $v = \lambda_0 + \lambda_0^{-1} > v_{\min}$. However, this behavior is still not the full story! A more complete analysis for the FKPP equation shows that for an initially sharp wavefront the velocity slowly approaches its asymptotic value as $v(t) = v_{\min}(1 - 3/Nt)$. Moreover, in the leading edge of the wave, the density is small and there is no reason that the evolution of the first few invaders can even be described by a continuum equation. If one accounts for the discrete nature of the particles, then the wave velocity is reduced compared to the continuum value: $v = v_{\min}(1 - \pi^2/\ln^2 N)$ — surprisingly rich behavior from such a basic equation of motion.

Conservative dynamics

In two-component alloy systems, the natural order parameter is the difference in the concentration of the two constituent elements. By its construction, this order parameter is conserved in an isolated piece of alloy. Thus dynamics different from the TDGL is needed to account for this microscopic conservation of material. At a phenomenological level, we seek a governing dynamical equation that ensures that the flux of atoms of each element of the alloy can be expressed in the form of a (conserved) continuity equation

$$\frac{\partial m}{\partial t} + \nabla \cdot \mathbf{J} = 0, \tag{8.6}$$

where we again consider overdamped dynamics so that the equation of motion is first order in time. In Eq. (8.6), the flux vector $\mathbf{J}$ should depend on the order parameter through the free energy (8.1). The

simplest choice that is both conservative and involves simple gradient flow is $\mathbf{J} \propto -\nabla \frac{\delta F}{\delta m}$. We again absorb the proportionality factor into the time scale to then obtain the equation of motion

$$\frac{\partial m}{\partial t} = \nabla^2 \frac{\delta F}{\delta m} = -\nabla^2 \left[\nabla^2 m - V'(m)\right]. \tag{8.7}$$

This is the *Cahn-Hilliard (CH) equation* for zero-temperature conservative dynamics.

The absence of thermal noise in Eqs. (8.2) & (8.7) suggests that we are effectively dealing with systems at zero temperature. To phenomenologically model the behavior at a positive temperature, we should add a (Langevin) noise term to the right-hand side of (8.2) or (8.7). There is a belief that this additional term should not change qualitative dynamical behavior as long as the temperature is below the critical temperature T_c. This belief boils down to the hypothesis that under renormalization, there are three possible behaviors:

- zero-temperature dynamics in which dynamical behavior is essentially the same for all $T < T_c$;
- critical dynamics ($T = T_c$);
- infinite-temperature dynamics in which the dynamical behavior for all $T > T_c$ are essentially the same, and trivial.

In this chapter, we will discuss zero-temperature dynamics because this case embodies many of the outstanding issues associated with coarsening. Typically, one starts with the initial temperature $T_i = \infty$ where the system is completely disordered and quenches the system to $T_f = 0$. The coarse-grained form of this disordered initial condition is

$$\langle m(\mathbf{x}, 0)\rangle = 0, \quad \langle m(\mathbf{x}, 0)\, m(\mathbf{x}', 0)\rangle = \delta(\mathbf{x} - \mathbf{x}'). \tag{8.8}$$

Thus the mathematical challenge is to determine the long-time behavior of the solutions of the *deterministic* nonlinear partial differential equations (8.2) and (8.7) subject to the *random* initial conditions (8.8).

8.2 Free Evolution

Numerical and analytical work has clearly shown that the solutions of (8.2) and (8.7) exhibit scaling, *i.e.*, a scale-invariant coarsening domain mosaic morphology develops at late times. This morphology is (statistically) independent of time when all lengths are rescaled by the typical domain size $L(t)$. This length scale grows algebraically with time, $L(t) \sim t^z$, with a nearly universal dynamical exponent z:

$$z = \begin{cases} 1/2 & \text{(TDGL)}; \\ 1/3 & \text{(CH)}; \end{cases} \tag{8.9}$$

for systems with a scalar order parameter. One important exception to Eq. (8.9) is one dimension, where we shall derive that domains grow logarithmically in time for the TDGL equation, $L(t) \sim \ln t$, so that the dynamical exponent $z = 0$.

As always, scaling greatly simplifies the description of the dynamics. The evidence in favor of scaling is compelling, but it has not been proven except for the TDGL in one dimension and for a small class of microscopic models, such as the one-dimensional Ising-Glauber model and the n-vector model with $n = \infty$. It is therefore important to show that scaling arises in the TDGL and the CH equations, even for an oversimplified setting. Such an example is provided by the toy model where the potential vanishes, $V(m) = 0$. The dynamical equation then reduces to the diffusion equation, $\frac{\partial m}{\partial t} = \nabla^2 m$, for a non-conserved order parameter, and to the bidiffusion equation, $\frac{\partial m}{\partial t} = -\nabla^4 m$, for a conserved order parameter. In these cases, the presence of a single growing length scale, $L(t) \sim t^{1/2}$ (TDGL), and $L(t) \sim t^{1/4}$ (CH), is obvious from dimensional analysis (see the subsections on dimensional analysis and scaling in Sec. 1.1 for further discussion of this point).

If a spin system is initially in the random state specified by Eq. (8.8), then the magnetization remains zero throughout the evolution. We therefore need two-body correlation functions to investigate the nature of

the coarsening; we already encountered this same issue in our study of the voter model and the Ising-Glauber model in the previous chapter. We thus define the two-body correlation function

$$C(\mathbf{r}_1,t_1,\mathbf{r}_2,t_2) \equiv= \frac{\langle m(\mathbf{r}_1,t_1)\, m(\mathbf{r}_2,t_2)\rangle}{\sqrt{\langle m^2(\mathbf{r}_1,t_1)\rangle\, \langle m^2(\mathbf{r}_2,t_2)\rangle}} \tag{8.10}$$

to probe the domain structure at two different space-time points. Our choice for the normalization makes the correlation function dimensionless. For simplicity, we use the shorthand $1 \equiv (\mathbf{x}_1,t_1)$ and $2 \equiv (\mathbf{x}_2,t_2)$. Translational invariance implies that $C(1,2) = C(\mathbf{r},t_1,t_2)$ with $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. It is also useful to explicitly study the *autocorrelation function* $A(t_1,t_2) = C(\mathbf{0},t_1,t_2)$ that measures the probability that the sign of the magnetization coincides at times t_1 and t_2.

Let's begin by investigating coarsening in the potential-free TDGL (diffusion) equation

$$\frac{\partial m}{\partial t} = \nabla^2 m \tag{8.11}$$

for the random initial condition of Eq. (8.8). The solution is

$$m(\mathbf{r}_1,t) = \frac{1}{(4\pi t)^{d/2}} \int m(\mathbf{z}_1,0)\, e^{-(\mathbf{r}_1-\mathbf{z}_1)^2/4t}\, d\mathbf{z}_1.$$

Then the average of the product of the magnetization at two different space-time points is

$$\begin{aligned}
\langle m(\mathbf{r}_1,t_1)m(\mathbf{r}_2,t_2)\rangle &= \frac{1}{[(4\pi)^2 t_1 t_2]^{d/2}} \iint \underbrace{\langle m(\mathbf{z}_1,0)\, m(\mathbf{z}_2,0)\rangle}_{\delta(\mathbf{z}_1-\mathbf{z}_2)} e^{-(\mathbf{r}_1-\mathbf{z}_1)^2/4t_1}\, e^{-(\mathbf{r}_2-\mathbf{z}_2)^2/4t_2}\; d\mathbf{z}_1 d\mathbf{z}_2 \\
&= \frac{1}{[(4\pi)^2 t_1 t_2]^{d/2}} \int e^{-(\mathbf{r}_1-\mathbf{z})^2/4t_1}\, e^{-(\mathbf{r}_2-\mathbf{z})^2/4t_2}\; d\mathbf{z} \\
&= \frac{1}{[(4\pi)^2 t_1 t_2]^{d/2}} \int e^{-\mathbf{r}_1^2/4t_1}\, e^{-\mathbf{r}_2^2/4t_2}\, e^{-z^2(1/4t_1+1/4t_2)}\, e^{\mathbf{z}\cdot(\mathbf{r}_1/2t_1+\mathbf{r}_2/2t_2)}\; .
\end{aligned}$$

We perform the integral in the last line by completing the square in the exponential to give

$$\langle m(\mathbf{r}_1,t_1)m(\mathbf{r}_2,t_2)\rangle = \frac{1}{[4\pi(t_1+t_2)]^{d/2}}\; e^{-(\mathbf{r}_1-\mathbf{r}_2)^2/4(t_1+t_2)}\; . \tag{8.12}$$

Finally, the normalized two-body correlation function defined by Eq. (8.10) is

$$C(1,2) = \left[\frac{2\sqrt{t_1 t_2}}{t_1+t_2}\right]^{d/4} e^{-(\mathbf{r}_1-\mathbf{r}_2)^2/4(t_1+t_2)}\; . \tag{8.13}$$

Let's now study coarsening for the potential-free CH (bidiffusion) equation

$$\frac{\partial m}{\partial t} = -\nabla^4 m\,. \tag{8.14}$$

Notice that the right-hand side now has a minus sign. The simplest way to solve the bidiffusion (and the diffusion) equations is with Fourier transforms. Using the Fourier expansion $m(\mathbf{r},t) = (2\pi)^{-d} \int \widehat{m}(\mathbf{k},t)\, e^{-i\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k}$ reduces (8.14) to an algebraic equation whose solution is simply $\widehat{m}(\mathbf{k},t) = \widehat{m}_0(\mathbf{k})\, e^{-k^4 t}$, with $\widehat{m}_0(\mathbf{k}) \equiv \widehat{m}(\mathbf{k},t=0)$. Inverting this Fourier transform gives

$$m(\mathbf{x},t) = (2\pi)^{-d} \int \widehat{m}_0(\mathbf{k})\, e^{-i\mathbf{k}\cdot\mathbf{x}-k^4 t}\, d\mathbf{k}. \tag{8.15}$$

For the white noise initial condition, the initial value of the correlation function in Fourier space is just

$$\langle \widehat{m}_0(\mathbf{k})\rangle = 0, \quad \langle \widehat{m}_0(\mathbf{k})\, \widehat{m}_0(\mathbf{k}')\rangle = (2\pi)^d \delta(\mathbf{k}+\mathbf{k}'). \tag{8.16}$$

Using these facts, the mean-square magnetization $\langle m^2(\mathbf{x},t)\rangle$ can be immediately found from Eqs. (8.8), (8.15), & (8.16), and the result is:

$$\langle m^2(\mathbf{x},t)\rangle = (2\pi)^{-d}\int e^{-2k^4t}\,d\mathbf{k} = B_d(2t)^{-d/4}\,, \tag{8.17}$$

where $B_d = (2\pi)^{-d}\Omega_d\Gamma(d/4)/4$, $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the area of the unit sphere in d dimensions, Γ is the Euler gamma function, and the subscript d refers to the spatial dimension of the system.

Similarly, the correlation function is

$$\langle m(1)\,m(2)\rangle = (2\pi)^{-d}\int e^{-i\mathbf{k}\cdot\mathbf{r}-k^4(t_1+t_2)}\,d\mathbf{k}\,. \tag{8.18}$$

When spatial points coincide, $\mathbf{r}=\mathbf{0}$, we obtain the autocorrelation function

$$A(t_1,t_2) = \left[\frac{2\sqrt{t_1t_2}}{t_1+t_2}\right]^{d/4}, \tag{8.19}$$

which has the same form for both the diffusion [Eq. (8.13)] and the bidiffusion equation. If the spatial points are different, $\mathbf{r}_1\neq\mathbf{r}_2$, we perform the integral in Eq. (8.18) by introducing spherical coordinates in $\mathbf{k}$-space for $d\geq 2$ so that $\mathbf{k}\cdot\mathbf{r} = kr\cos\theta$ and $d\mathbf{k} = k^{d-1}\,\Omega_{d-1}\sin^{d-2}\theta\,d\theta\,dk$ to give

$$\langle m(1)\,m(2)\rangle = \frac{\Omega_{d-1}}{(2\pi)^d}\int_0^\infty k^{d-1}\,e^{-k^4(t_1+t_2)}\,F_d(kr)\,dk\,,$$

where

$$F_d(u) = \int_0^\pi (\sin\theta)^{d-2}\,e^{-iu\cos\theta}\,d\theta\,.$$

This integral may be readily determined for $d\leq 3$ and the two-body correlation function $C(r,t_1,t_2)$ is:

$$C(r,t_1,t_2) = \begin{cases} \dfrac{2(4t_1t_2)^{1/8}}{\Gamma(1/4)\,r}\displaystyle\int_{-\infty}^{\infty}\cos q\;e^{-q^4\tau}\,dq & d=1;\\[2ex] \dfrac{4(4t_1t_2)^{1/4}}{\sqrt{\pi}\,r^2}\displaystyle\int_0^\infty q\,I_0(q)\;e^{-q^4\tau}\,dq & d=2;\\[2ex] \dfrac{2(4t_1t_2)^{3/8}}{\Gamma(3/4)\,r^3}\displaystyle\int_0^\infty q\,\sin q\;e^{-q^4\tau}\,dq & d=3. \end{cases} \tag{8.20}$$

In these equations, we use shorthand notations $q=kr$, $\tau=(t_1+t_2)/r^4$, and I_0 is the modified Bessel function of order 0.

Finally, the *temporally normalized* correlation function $G(\mathbf{r},t_1,t_2)\equiv C(\mathbf{r},t_1,t_2)/A(t_1,t_2)$ is a function of a *single* scaling variable τ (here the term "temporally normalized" reflects the property $G(\mathbf{0},t_1,t_2)\equiv 1$). This single-parameter scaling is a peculiarity of the toy model rather than a universal rule; generally, rotational symmetry and dynamical scaling would imply that $G(\mathbf{r},t_1,t_2)$ is a function of two variables. Finally at equal times $t_1=t_2=t$, both correlation functions $C(\mathbf{r},t,t)$ and $G(\mathbf{r},t,t)$ reduce to the equal-time correlation function $G(\tau)$ which is the function of the single scaling variable $\tau = 2t/r^4$. For the potential-free CH equation, the precise form of the equal-time correlation function $G(\tau)$ is

$$G(\tau) = \begin{cases} \dfrac{2\tau^{1/4}}{\Gamma(1/4)}\displaystyle\int_{-\infty}^{\infty}\cos q\;e^{-q^4\tau}\,dq & d=1;\\[2ex] \dfrac{4\tau^{1/2}}{\sqrt{\pi}}\displaystyle\int_0^\infty q\,I_0(q)\;e^{-q^4\tau}\,dq & d=2;\\[2ex] \dfrac{2\tau^{3/4}}{\Gamma(3/4)}\displaystyle\int_0^\infty q\,\sin q\;e^{-q^4\tau}\,dq & d=3. \end{cases}$$

8.3 Case Studies in Non-Conservative Dynamics

We now turn to explicit solutions of the basic equations of motion. Because domain interfaces in coarsening are geometrically complex, we first study the simplest settings that admit an exact analysis yet still allow for non-trivial dynamics. There are several classical such examples, such as a single domain wall and a single spherical droplet. We also discuss additional examples that are not so widely appreciated yet are rich laboratories to develop understanding. These include a single finger, a single wedge, and an isolated shrinking grain.

Straight domain wall

The simplest example of an interface is a straight domain wall. For the discrete Ising-Glauber model, this flat interface is trivial because it is merely a step function. For the continuum TDGL equation, a flat interface is a bit more interesting because the coarse-grained magnetization has a non-trivial spatial variation across the interface. Since a flat interface is stationary, the TDGL equation reduces to the ordinary differential equation for the order parameter

$$\frac{d^2m}{dx^2} = V'(m), \tag{8.21}$$

subject to the boundary conditions $m(\pm\infty) = \pm 1$ that lead to the domain wall being stationary.

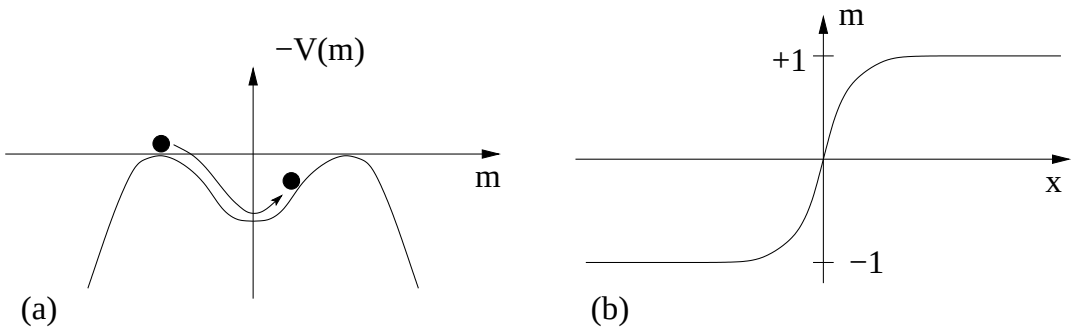


Figure 8.1: (a) The effective potential $-V(x) = -\frac{1}{2}(1-m^2)^2$ and the motion of the analogue particle that corresponds to a single kink. The particle starts at the top of the left maximum at $m = -1$ and moves without dissipation until it eventually stops at the maximum at $m = +1$. (b) The corresponding dependence of m on x.

Before solving (8.21), we emphasize a neat interpretation of the domain wall in terms of classical mechanics: Treating x as the time variable and m as the coordinate, we see that the problem is equivalent to the motion of a fictitious, analogue unit-mass particle in the potential $-V$. The energy of this analogue particle

$$E = \frac{1}{2}\left(\frac{dm}{dx}\right)^2 - V(m) \tag{8.22}$$

does not change with "time" x. The boundary conditions imply that the total energy equals zero. Thus $\frac{dm}{dx} = \sqrt{2V(m)}$, or

$$x = \int_{m(0)}^{m(x)} \frac{d\mu}{\sqrt{2V(\mu)}}\,, \tag{8.23}$$

where it is customary to set the "initial condition" to be $m(x=0) = 0$; this then corresponds to the "energy" $E = 0$. For the classic potential $V(m) = \frac{1}{2}\,(1-m^2)^2$, the domain wall profile from solving (8.23) is then

$$m(x) = \tanh x\,. \tag{8.24}$$

This solution is known as a "kink" in a ϕ^4 theory. The width of this kink is of the order of 1, while its free energy density is

$$\frac{1}{2}\left(\frac{dm}{dx}\right)^2 + V(m) = 2V(m(x)) = \frac{1}{\cosh^4 x}\,, \tag{8.25}$$

and the total free energy per unit area, that is, the surface tension, is

$$\sigma = \int_{-\infty}^{\infty} 2V(m)\,dx = \int_{-\infty}^{\infty} 2V(m)\,\left(\frac{dm}{dx}\right)^{-1} dm = \int_{-1}^{1} \sqrt{2V(m)}\,dm = \frac{4}{3}\,. \tag{8.26}$$

Within the classical mechanical analogy, the governing equation (8.21) can be derived by minimizing the "action"

$$\int_{-\infty}^{\infty} \left[\frac{1}{2}\left(\frac{dm}{dx}\right)^2 + V(m)\right] dx,$$

and because the kinetic and potential energies are equal, the action associated with the analogue particle coincides with the surface tension.

The mechanical analogy provides many useful insights about the phase behavior of a system in situations where an analytical description of domain walls is not feasible. For example, suppose that the initial condition is a combination of two static solutions, such as a kink-antikink configuration. That is, the order parameter starts at $m \approx -1$ for $x \to \infty$ and increases to $m \approx +1$ over some intermediate range of x, and then returns to the value $m \approx -1$ for $x \to +\infty$. Can such a configuration be stable? Within the mechanical analogy, we are asking if it is possible for the analogue particle to start at the maximum at $m = -1$, move to the maximum at $m = +1$, and then return to the maximum at $m = -1$ (see Fig. 8.1). Clearly this motion is impossible, so we conclude without calculation that a kink-antikink pair cannot be stable. Numerical simulations show that the kink and antikink slowly move toward each other and eventually annihilate. We shall analyze this motion later in this chapter.

Another useful aspect of the mechanical analogy is that it can characterize *all* possible one-dimensional stationary solutions that correspond to a finite free energy per unit area. For any potential function, the analogue particle (i) must start at one maximum of the potential $[-V]$ at "time" $x = -\infty$ and reach another maximum of the potential $[-V]$ at $x = \infty$, and (ii) if the maxima are different, they must be adjacent. Thus for potentials with two degenerate minima (at $m = \pm 1$) there are two types of static solutions — a kink and an antikink — that we symbolically denote as $[-1, 1]$ and $[1, -1]$, respectively. Similarly for potentials with three three degenerate minima (for example, a 3-state spin system with minima at $m = \pm 1$ and $m = 0$) there are four types of static solutions: $[-1, 0], [0, -1], [1, 0], [0, 1]$. A domain wall of the form $[-1, 1]$ cannot occur. To be concrete, for the potential $V(m) = \frac{1}{2}\,m^2\,(1 - m^2)^2$, the exact static solutions of Eq. (8.23) that separate the phases with $m = 0$ and $m = \pm 1$ are

$$m(x) = \pm\left(1 + e^{\pm x}\right)^{-1/2}\,. \tag{8.27}$$

Higher dimensions

Perhaps the simplest interface dynamics example in higher dimension is the evolution of a small droplet of one phase with $m = -1$ in a sea of the opposite phase with $m = +1$. From the perspective of individual spins, we expect that a small cluster of minority spins will be overwhelmed by the large majority of oppositely-oriented spins. This is precisely what happens in the continuum picture. Moreover, the consideration of a small droplet avoids all the complexities associated with the tortuous interface between two competing phases, as the droplet becomes spherical as it shrinks to zero at long times.

Since the droplet shrinks, we must use the full time-dependent TDGL equation to describe its evolution. Due to spherical symmetry, the TDGL equation (8.2) simplifies to:

$$\frac{\partial m}{\partial t} = \frac{\partial^2 m}{\partial r^2} + \frac{d-1}{r}\,\frac{\partial m}{\partial r} - V'(m). \tag{8.28}$$

Provided that the droplet radius R greatly exceeds the interface width, we anticipate that the magnetization has the form $m(r, t) = f(r - R(t))$. Substituting this ansatz into (8.28) and defining $x = r - R(t)$ gives

$$f'' + \left(\frac{d-1}{r} + \frac{dR}{dt}\right) f' - V'(f) = 0. \tag{8.29}$$

Let's now multiply this equation by f' and integrate with respect to x through the interface. For a large droplet, the interface is relatively sharp and localized to a narrow region around a value $R(t)$. Consequently, we can set the lower limit of the integral, $-R$, to $-\infty$. This gives

$$(f')^2\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(\frac{d-1}{r} + \frac{dR}{dt}\right)(f')^2\, dx - V(x)\Big|_{-\infty}^{\infty} = 0. \tag{8.30}$$

We now use the fact that $f(x)$ changes suddenly from -1 to 1 near $x = 0$ (corresponding to $r = R(t)$), so that f' is sharply peaked near the origin. Thus the first term vanishes. Similarly, the last term is zero, since the potential has the same value at $x = \pm\infty$. Finally, the integral is non-zero only very close to $x = 0$ and we may replace the integrand by its value at $r = R$ to give

$$\frac{d-1}{R} + \frac{dR}{dt} = 0. \tag{8.31}$$

Solving this equation gives the fundamental result that a small droplet shrinks as $R^2 = R_0^2 - 2(d-1)t$, so that the time τ for a droplet to disappear is $\tau \propto R_0^2$.

The virtue of this focus on the interfacial region is that, instead of the unsolvable TDGL equation (8.28), we obtain a more tractable equation for the interface dynamics. This simplified description of interface dynamics applies for an arbitrarily complex domain geometry, provided that the characteristic length scale associated with the interface morphology greatly exceeds the interface width. This construction is particularly simple in two dimensions. Locally an interface has a curvature $1/R$ and therefore the interface can be approximated by a circle of radius R. As a result, the velocity normal to the interface is $v \approx -1/R$. Generally in d dimensions, the interface is a $(d-1)$-dimensional manifold and we denote by $R_1, \ldots, R_{d-1}$ its principal radii of curvature (which depend on the local position on the interface). Then the normal interface velocity is the sum of the principal curvatures $1/R_j$; this leads to the *Allen-Cahn* (AC) equation for this velocity:

$$v = -(d-1)K, \tag{8.32}$$

where K is the *mean curvature*, $K = (d-1)^{-1}\sum 1/R_j$.

There are two noteworthy points about the AC equation. First, the AC equation does not apply in one dimension, where the interface between two domains is a point. Here the interface velocity is an exponentially small function of the domain lengths on either side of the interface. This feature leads to domains that grow logarithmically in time and to extremal dynamics in which only the smallest domain merges with its neighbors in an update event. A second, more important, point is that the interface dynamics derived from the TDGL equation is an example of *curvature-driven flow*, in which general shapes evolve in arbitrary dimension due to a local velocity that is proportional to the local curvature. Curvature-driven flows have been thoroughly investigated and much is known about evolution of a single closed interface. In two dimensions, every such interface asymptotically approaches a (shrinking) circular shape. The same happens in three dimensions if the initial interface is everywhere *convex*. However, an interface that contains both convex and concave portions can fission if the concave portion of the manifold is sufficiently thin, such as an elongated liquid drop with an extremely thin neck of fluid in the middle. The classification of all possible topology changes in higher dimensions due to curvature-driven flow is still incomplete.

Two-dimensional unbounded domains

A beautiful set of domain evolution problems is inspired by returning to a discrete Ising spin system on the square lattice and asking: what are the simplest initial states for which evolution under zero-temperature Glauber dynamics can be studied analytically? A example that is too simple because it is static is a single straight domain wall that cuts a finite system in two. To have a system that actually evolves, the initial interface must contain at least one corner and we now study two such examples (Fig. 8.2): (a) the semi-infinite finger (2 initial corners), and (b) the 90° infinite wedge (1 initial corner). Curiously, although the finger appears to be more complex than the wedge because there are two initial corners, its dynamics is simpler than that of the wedge. The wedge is, in fact, extremely rich, with nice connections between the shape of the wedge and the famous number theoretic-problem of partitioning of the integers.

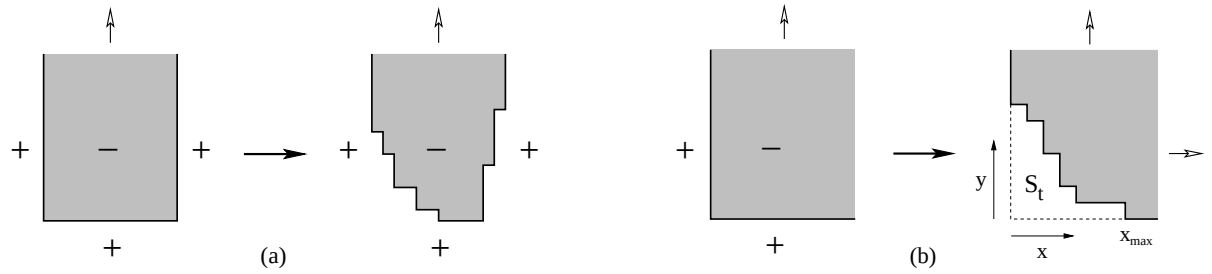


Figure 8.2: (a) A semi-infinite finger and (b) a 90° infinite wedge, showing both the initial state (left of each figure) and the system at a later time (right side). For the wedge, the evolving interface encloses an area S_t at time t.

Let's begin with the semi-infinite finger. At the individual spin level, the interface develops a series of terraces that culminate at the tip. At the tip isolated islands can even pinch off. A detailed understanding of these microscopic details has yet to be achieved. However, within the continuum description the asymptotic behavior is quite simple. Let the minority phase initially occupy the semi-infinite region $y > 0$ and $|x| < L$. The interesting limit is $t \gg L^2$ where sufficient time has elapsed so that the two corners of the initial finger interact. In this long-time regime, the finger relaxes to a limiting shape that recedes at a constant velocity. In a reference frame that moves with the finger, the interface $y(x)$ is then stationary.

For this geometry, the AC equation is

$$v_n = \frac{y''}{[1+(y')^2]^{3/2}} \ . \tag{8.33}$$

Here v_n is the velocity normal to the interface. The right-hand side is just the curvature of the locus $y(x)$ that defines the finger boundary and the prime denotes differentiation with respect to x. In the steady state, only the vertical component of the velocity, given by $v_y = v_n/\cos\theta = v_n\sqrt{1+(y')^2}$, is relevant (see Fig. 8.3). Thus the vertical velocity satisfies the equation

$$v_y = \frac{y''}{1+(y')^2} \ , \tag{8.34}$$

subject to the boundary condition $y \to \infty$ when $|x| \to L$. The solution to this equation is

$$y = -\frac{2L}{\pi} \ln\left[\cos\left(\frac{\pi x}{2L}\right)\right] \ . \tag{8.35}$$

Substituting this profile into Eq. (8.34), we find that the finger recedes at a constant velocity $v = \pi/2L$.

Next we study the infinite wedge (Fig. 8.2). For concreteness, let the corner be at the origin so that the wedge initially occupies the region $x, y \geq 0$. The corner of the wedge recedes diffusively, $x, y \propto \sqrt{t}$, both in Ising-Glauber dynamics and in the continuum description. Because of the absence of any constant with dimension of length in Eq. (8.34), the corresponding solution admits the self-similar form

$$X = x(t)/\sqrt{t}, \qquad Y(X) = y(x,t)/\sqrt{t}\,, \quad . \tag{8.36}$$

Notice that the increase of the magnetization is equal to twice the area under the curve $y(x,t)$. From (8.36), the area is proportional to t, so that the magnetization also grows linearly with time. To determine the evolution of the wedge, we substitute the ansatz Eq. (8.36) into the equation of motion (8.34) and find that the scaling function $Y(X)$ obeys

$$\frac{Y - XY'}{2} = \frac{Y''}{1+(Y')^2}, \tag{8.37}$$

where prime indicates differentiation with respect to X. Equation (8.37) should be solved subject to the constraints $\lim_{X\to\infty} Y(X) = 0$ and $\lim_{X\to+0} Y(X) = \infty$ that correspond to the wedge geometry. A wedge that initially occupies the region $y > |x|\tan\theta$ could also be studied if we impose the constraints $Y \to \pm X\tan\theta$ as $X \to \pm\infty$.

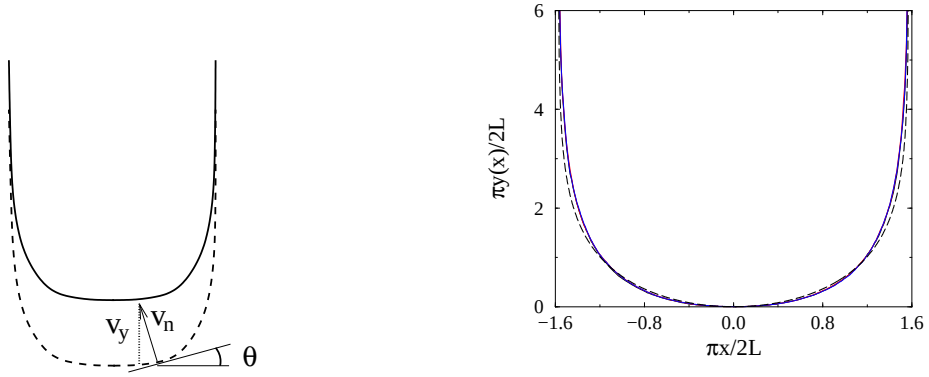


Figure 8.3: (Left) Schematic illustration of the receding finger, showing the normal velocity v_n and the vertical velocity v_y. (Right) Comparison of the finger shape predicted by the TDGL approach given in Eq. 8.35 (dashed) with simulation results for a system of width $2L = 400$ at long times (solid).

We introduce the polar coordinates $(X, Y) = (r\cos\theta, r\sin\theta)$ to recast Eq. (8.37) into the following equation for $r = r(\theta)$:

$$2r\,\frac{d^2r}{d\theta^2} - \left(4 + r^2\right)\left(\frac{dr}{d\theta}\right)^2 = r^2\left(2 + r^2\right). \tag{8.38}$$

Writing $\frac{dr}{d\theta} = R(r)$, reduces Eq. (8.38) to a first-order equation whose solution is $R^2 = r^4\, e^{r^2/2}\, F(r, r_0)$, with

$$F(r, r_0) = \int_{r_0}^{r} \left(\frac{2}{\rho^3} + \frac{1}{\rho}\right) e^{-\rho^2/2}\, d\rho\,, \tag{8.39}$$

and r_0 is the scaled distance from the origin to the closest point on the interface. The interface is now determined from

$$\frac{dr}{d\theta} = \pm r^2\, e^{r^2/4}\sqrt{F(r, r_0)}, \tag{8.40}$$

with the plus sign for $\theta \geq \pi/4$ and the minus sign for $\theta \leq \pi/4$ to give a symmetric solution about the diagonal $\theta < \pi/4$. Integrating Eq. (8.40) gives the explicit equation for $\theta = \theta(r)$

$$\theta = \int_{r}^{\infty} \rho^{-2}\, e^{-\rho^2/4}\, [F(\rho, r_0)]^{-1/2}\, d\rho \tag{8.41}$$

for $\theta \leq \pi/4$. For $\pi/4 < \theta < \pi/2$, we simply use $r(\theta) = r(\frac{\pi}{2} - \theta)$ to ensure symmetry of the interface about the diagonal. The value of the unknown r_0 may now be obtained by ensuring that $\theta = \pi/4$ when $r = r_0$. This gives the criterion

$$\int_{r_0}^{\infty} r^{-2}\, e^{-r^2/4} [F(r, r_0)]^{-1/2} = \frac{\pi}{4}\, dr\,, \tag{8.42}$$

whose numerical solution is $r_0 \approx 1.0445$. Equation (8.41), with F given by (8.39), provides an explicit representation of $\theta(r)$ on the interface in terms of the (scaled) distance $r \in [r_0, \infty)$ from the origin. For $r \to \infty$, the interface becomes much simpler. From Eqs. (8.39) and (8.41) we obtain $\theta \to A\, r^{-3}\, e^{-r^2/4}$ with $A = 2\,[F(\infty, r_0)]^{-1/2}$ which, in Cartesian coordinates, is

$$Y \to A\, X^{-2}\, e^{-X^2/4}\,. \tag{8.43}$$

Apart from the numerical factor A, the asymptotic behavior follows directly from Eq. (8.37) after dropping the subdominant terms.

The staircase profile of the interface in Fig. 8.2(b) has a simple relation to the famous number-theoretic problem of the partitions of the integers. Given an arbitrary integer N, how many different ways can we break up this number into k integer pieces $n_1, n_2, \ldots, n_k$ with $\sum_k n_k = N$?. Geometrically, this question is

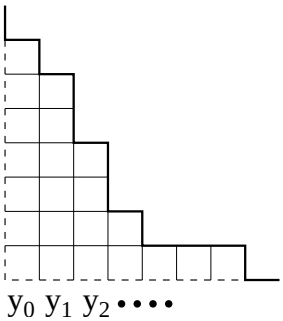


Figure 8.4: The Young diagram that is based on the interface profile of Fig. 8.2. This diagram corresponds to a partition of the integer 22 into the set $\{7, 6, 4, 2, 1, 1, 1\}$.

the same as enumerating all possible interfaces with a given number of flipped spins in the wedge geometry! The interface height y_k and the k^{th} column is same as n_k in the partitioning problem. For $N \to \infty$, the staircase generated by its partition approaches a limiting shape. The equation for this shape is remarkably simple:

$$e^{-\lambda x} + e^{-\lambda y} = 1, \quad \text{with } \lambda = \frac{\pi}{6\sqrt{N}}. \tag{8.44}$$

In scaled units, the asymptotic shape of the integer partition limiting staircase is $Y \sim \frac{\sqrt{6}}{\pi} e^{-\pi X/\sqrt{6}}$, whereas the asymptotic Ising staircase is given by $Y \sim \frac{1.0445\ldots}{X^2}\, e^{-X^2/4}$. The existence of two distinct shapes arises from the different construction rules of the Ising interface and the interface generated by the integer partitions. In the partitioning of the integers, each staircase corresponding to fixed N is generated with the same weight. On the other hand, the Ising interface is generated dynamically and when its area reaches N, the interface contains an imprint of its entire history.

8.4 Conservative Dynamics

The influence of a conservation law severely limits the way in which an interface between two domains can move. While an interface should move to reduce its local curvature and consequently the energy, such an evolution has to be accompanied by a global rearrangement of interfaces to ensure that the order parameter is conserved. While the full problem remains quite open, much progress has again been made in studying the simplest geometries, such as a single droplet or a dilute population of droplets immersed in a majority phase. Here, each droplet remains spherical throughout the evolution, so that one avoids the geometrical complexity associated with random interfaces. Nevertheless, the kinetics of this dilute limit is still quite subtle and rich.

Evolution of a single droplet

As an essential preliminary, consider the evolution of a single droplet of one phase that is immersed in a background of the opposite phase. This statement, however, has a basic inconsistency that requires clarification. If the two phases were perfectly separated, then the droplet could not evolve because any change in the droplet size would violate the conservation of the order parameter. Instead, we should think of a single drop of liquid that is floating in a closed container of gas that is saturated with the liquid in the vapor phase. At the surface of the droplet there is both evaporation as well as condensation of vapor molecules back into the droplet. It is through the combination of these two processes that the droplet can evolve while still conserving the order parameters.

The condensation rate is given by the flux of vapor molecules to the droplet which, in turn, is determined by vapor concentration in the gas. In principle, the vapor concentration exterior to the droplet obeys the diffusion equation. However, because the droplet radius changes slowly with time, we apply the quasi-static approximation (see Sec. 2.5) in which we ignore the time dependence and deal with the much simpler

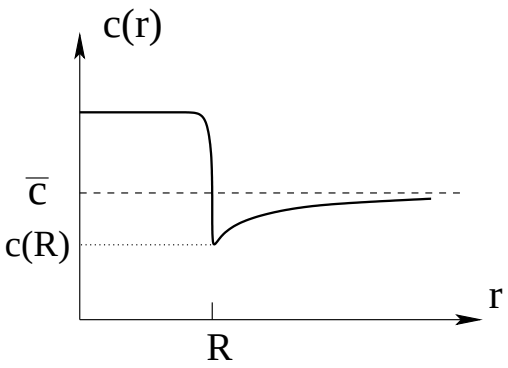


Figure 8.5: Dependence of the concentration of the minority phase as a function of radius.

Laplace equation to determine the external vapor concentration. Schematically, the concentration of the minority phase as a function of distance from the center of a droplet of radius R should then have the dependence sketched in Fig. 8.5. The average concentration of the minority phase in the entire system $\bar{c}$ must be larger than the supersaturation value c_∞, *i.e.*, there must be supersaturation, so that a droplet can form in the first place. (Conversely, for $\bar{c} < c_\infty$, the minority phase remains as a homogeneous vapor in the gas.) Inside the droplet, the concentration is much higher than $\bar{c}$ by the very definition of a droplet. Outside the droplet, the vapor concentration obeys Laplace's equation with the boundary conditions $c(r \to \infty) \to \bar{c}$ and $c(R)$ determined by the Gibbs-Thompson relation (see below). In three dimensions, $c(r)$ therefore has the dependence

$$c(r) = \bar{c} - [\bar{c} - c(R)]\,\frac{R}{r}\ . \tag{8.45}$$

The Gibbs-Thompson relation relates $c(R)$ to c_∞ by the following physical picture: for a bulk liquid in equilibrium with a saturated vapor, the vapor concentration at the interface must be c_∞ by definition. For a small droplet, the vapor concentration is larger than that of a bulk liquid because a larger fraction of molecules are at the interface. Thus $c(R)$ should exceed c_∞ by an amount that vanishes as the droplet radius goes to infinity. This relationship is encapsulated by the Gibbs-Thompson relation $c(R) = c_\infty\left(1 + \frac{\nu}{R}\right)$, where ν is known as the capillary length and is simply related to the surface tension and the temperature.

Assembling these elements, the volume V of an isolated droplet changes with rate

$$\frac{dV}{dt} = 4\pi R^2\ \frac{dR}{dt} = 4\pi R^2\, D\frac{\partial c}{\partial r}\Big|_{r=R}\ , \tag{8.46}$$

from which $\dot{R} = D[\bar{c} - c(R)]/R$. We now define $\Delta = \bar{c} - c_\infty$ as the degree of supersaturation, and use the Gibbs-Thompson relation to eliminate $c(R)$ in (8.46) to give

$$\frac{dR}{dt} = \frac{D}{R}\left(\Delta - c_\infty\frac{\nu}{R}\right) \equiv \frac{\alpha}{R^2}\left(\frac{R}{R_c} - 1\right). \tag{8.47}$$

From the latter form, it is clear that for $R > R_c = c_\infty\nu/\Delta$ the droplet grows, while for $R < R_c$ the droplet shrinks. More importantly, we see by power counting that this equation contains the seeds of $t^{1/3}$ coarsening. An isolated supercritical droplet asymptotically grows as $t^{1/3}$, while a subcritical droplet shrinks to zero size in a time that scales as the cube root of its initial radius.

Lifshitz-Slyazov-Wagner coarsening

The foregoing discussion provides the basis for understanding the coarsening of a dilute heterogeneous population of droplets in a supersaturated background. This problem was independently analyzed by Lifshitz and Slyozov, and by Wagner, and we refer to the problem as LSW coarsening. The dilute limit allows us to make the approximation that droplets are non-interacting so that the concentration field around each droplet is the same as that of an isolated droplet. The basic feature of this coarsening is already contained in Eq. (8.47): droplets whose radii is larger than R_c grow and those with smaller radii shrink. In a population

of heterogeneous droplets, the value of R_c has to be determined self consistently, and we show that this calculation gives $R_c \propto t^{1/3}$.

Let $f(R,t)$, be the concentration of droplets of radius R at time t. This concentration evolves by the continuity equation

$$\frac{\partial f}{\partial t} + \frac{\partial j}{\partial R} = 0, \tag{8.48}$$

where the flux $j = \dot{R} f(R,t)$ is just the difference between the increase and decrease of droplets of radius R due to their evolution. We wish to solve this equation of motion with $\dot{R}$ given by Eq. (8.47) and subject to the constraint that the total mass of the minority phase is conserved. This constraint may be written as

$$\bar{c} - c_\infty + \frac{4\pi}{3}\int_0^\infty R^3\, f(R,t)\, dR = \text{const.} \tag{8.49}$$

In the minority limit, the volume fraction of the minority phase that exists as freely diffusing monomers is vanishingly small. Thus the conservation law reduces to the condition that the total volume of the droplets is fixed. With this proviso, Eqs. (8.47)–(8.49) constitute the governing equations of coarsening in the minority limit with a conserved order parameter.

To solve these equations of coarsening, it is again very useful to apply scaling. The natural scaling ansatz for this system, under the assumption that the mass of all droplets is conserved, is

$$f(R,t) = \frac{1}{R_c^4}\,\phi\left(\frac{R}{R_c}\right).$$

Here the prefactor is determined by the conservation of the total mass of the minority phase, namely, $\int R^3\, f(R,t)\, dR = \text{const}$. Substituting this scaling ansatz into (8.48), the first term in this equation becomes

$$\frac{\partial f}{\partial t} = -\frac{4\dot{R}_c}{R_c^5}\,\phi - \frac{1}{R_c^4}\,\phi'\,\frac{R\dot{R}_c}{R_c^2} = -\frac{\dot{R}_c}{R_c^5}(4\phi + z\phi'),$$

where the prime denotes differentiation with respect to the scaled variable $z \equiv R/R_c$. Similarly, the second term in (8.48) becomes

$$\frac{\partial}{\partial R}\left[\frac{\alpha}{R^2}\left(\frac{R}{R_c} - 1\right)\frac{1}{R_c^4}\,\phi\right] = \frac{\alpha}{R_c^7}\left[\left(\frac{1}{z} - \frac{1}{z^2}\right)\phi' + \left(\frac{2}{z^3} - \frac{1}{z^2}\right)\phi\right].$$

With these preliminaries the partial differential equation of (8.48) is converted into the ordinary differential equation

$$R_c^2\dot{R}_c = \alpha\,\frac{\left[\left(\frac{1}{z} - \frac{1}{z^2}\right)\phi' + \left(\frac{2}{z^3} - \frac{1}{z^2}\right)\phi\right]}{4\phi + z\phi'} \equiv \alpha\gamma, \tag{8.50}$$

where we rearranged the terms to put all the time dependence of the left and all the z dependence on the right. Since both sides are functions of different variables they each must be constant. In the simplest applications of scaling, such as the scaling solution to the diffusion equation in Sec. 1.1, the value of the separation constant plays little role in the scaling solution. In contrast, for coarsening, the separation constant γ is essential. Depending on the value of γ, there are three different regimes of behavior, only one of which is physically meaningful.

Let us now examine the separated equations and thus determine the condition that gives the physical value of γ. For the time dependence we have

$$R_c^2\,\dot{R}_c = \alpha\gamma, \tag{8.51}$$

with solution

$$R_c(t) = (3\alpha\gamma t)^{1/3}. \tag{8.52}$$

Thus coarsening under the constraint of a conserved order parameter leads to a $t^{1/3}$ growth of the typical droplet radius. This slower-than-diffusive growth of the typical droplet also justifies for the quasi-static

approximation that was used to determine the concentration outside a droplet. For the z dependence of the scaling function, the governing equation (8.50) becomes, after some simple rearrangement:

$$a\phi' + b\phi = 0, \qquad \text{with} \qquad a = -\frac{1}{z^2} + \frac{1}{z} - z\gamma, \quad b = \frac{2}{z^3} - \frac{1}{z^2} - 4\gamma\,. \tag{8.53}$$

Thus the scaling function is formally given by

$$\ln\phi = \int^z -\frac{b(y)}{a(y)}\,dy = \int^z \frac{2 - y - 4\gamma y^3}{1 - y + \gamma y^3}\,\frac{dy}{y}\,. \tag{8.54}$$

Thus far, the reasoning is standard: we've used scaling to separate the partial differential equation (8.48) into two ordinary differential equations. As is generally the case, the time dependence then follows easily. For LSW coarsening, the analysis of the z dependence is subtle because of the essential role of the separation constant γ. The first basic and somewhat surprising consequence of Eq. (8.54) is that $\phi(z)$ must have a sharp cutoff at a value $z_{\max}$, beyond which $\phi(z) = 0$. To demonstrate this fact, suppose the opposite is true. Then as $z \to \infty$, Eq. (8.54) would tell us that $\phi(z)$ should asymptotically vary as

$$\ln\phi \sim \int^z -4\,\frac{dy}{y} \sim -4\,\ln z \longrightarrow \phi \sim z^{-4}\,.$$

A power law tail for ϕ is impossible, however, as this asymptotic decay would lead to a divergence of the total mass of the minority phase:

$$\int R^3\,\phi\,(R,t)\,dR \sim \int z^3\,z^{-4}\,dz \to \infty.$$

Thus we conclude that $\phi(z)$ has a sharp cutoff beyond some value $z_{\max}$.

A second and more profound fact is that only one value of the separation constant γ is physically allowed. To see why this is the case, let's re-examine the behavior of $\dot{R}$ in scaled units. Using Eqs. (8.47) and (8.51), we find, after some simple algebra:

$$\dot{z} = \frac{1}{3\gamma t}\left(\frac{1}{z} - \frac{1}{z^2} - \gamma z\right) \equiv \frac{1}{3\gamma t}\;g(z)\,. \tag{8.55}$$

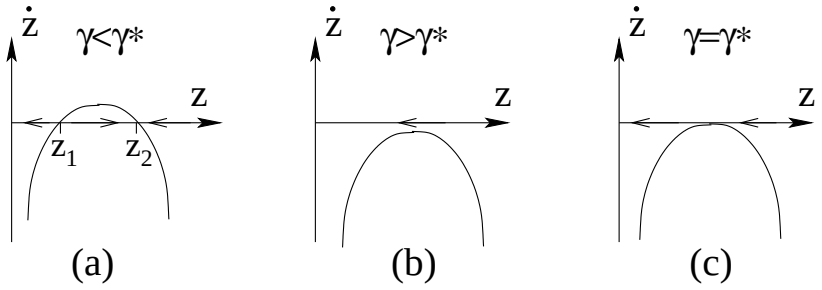


Figure 8.6: Sketch of $\dot{z} = g(z)/(3\gamma t)$ versus z for the 3 cases: (a) $\gamma < \gamma^*$, (b) $\gamma > \gamma^*$, and (c) $\gamma = \gamma^*$. The arrows on the z-axis shows the flow of z in Eq. (8.55).

Let's now examine the behavior of $g(z)$ for different values of γ. There are 3 cases:

(a) $\gamma < \gamma^* = 4/27$. Here $z(t)$ flows exponentially quickly to the stable fixed point at z_2 for $z(0) > z_1$ (Fig. 8.6). That is, the radii of all droplets approach the common value $R_c z_2$, which diverges as $t \to \infty$. Such a distribution cannot satisfy mass conservation.

(b) $\gamma > \gamma*$. Here $z(t) \to 0$ exponentially quickly in time for any initial value of z. Thus all droplets shrink to zero and mass conservation is again violated.

(c) $\gamma = \gamma^*$. In this case $z = 3/2$ is a fixed point, but one that is approached as a power law in time. This subtle behavior is the mechanism that allows mass conservation to be satisfied. If the fixed point at $z = 3/2$ was reached exponentially quickly, then again all droplet radii would approach the common value of $3R_c/2$ and mass conservation would be violated. The slow decrease in z ensures the delicate balance between growth and shrinking of clusters in a mass-conserving way.

For the physical case of $\gamma = \gamma^*$, Eq. (8.54) for the scaling function can be factorized as

$$\ln \phi = \int^z \frac{2 - y - \frac{16}{27}y^3}{(y - \frac{3}{3})^2(y+3)} \, \frac{dy}{y} \; .$$

Evaluating the latter integral by a partial fraction expansion now gives the quite complex form of the scaling function (in three dimensions):

$$\begin{aligned} \phi(z) &\propto z^2 \, (z+3)^{-7/3} \, (3-2z)^{-11/3} \, e^{-3/(3-2z)} \quad && z < 3/2; \\ &= 0 && z > 3/2, \end{aligned} \tag{8.56}$$

and a plot of this scaled droplet radius distribution is shown in Fig. 8.7.

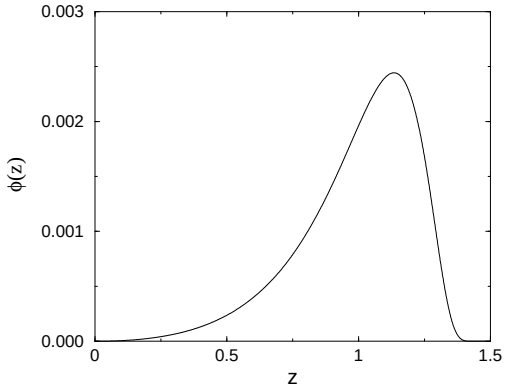


Figure 8.7: The scaling function $\phi(z)$ for LSW coarsening given by Eq. (8.56).

8.5 Extremal Dynamics

Shrinking of a single domain

While power-law domain growth is a generic feature of zero-temperature coarsening, there is one important case where much slower logarithmic growth occurs — the TDGL equation in one dimension. The case of one dimension is special because there is no local curvature to drive an interface; rather the interface moves by a net flux of order parameter across neighboring domains that is determined, in turn, by their lengths. In the long-time limit, this flux vanishes as e^{-L}, where L is the length of a typical domain; the smallness of this flux is responsible for logarithmic coarsening. As a consequence of this slow growth, the asymptotic dynamics of the one-dimensional TDGL equation has an extremal character in which only the smallest domain in the system merges with its two nearest neighbors in a single update state. This extremal picture then provides a natural way to determine the domain dynamics.

To understand the origin of the slow domain growth, consider a single large domain of length $L = x_2 - x_1$ of magnetization is close to $+1$ that is immersed in a sea with the magnetization close to -1 (Fig. 8.8). The two interfaces of this bubble consist of a kink at $x = x_1(t)$ and an antikink at $x = x_2(t) = x_1 + L$. If L is much larger than the width of each interface, which is a good approximation at long times, the kink and

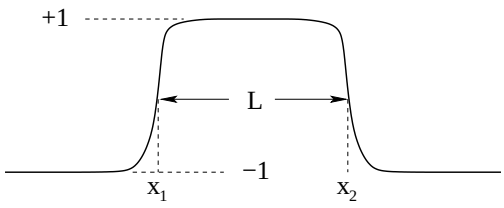


Figure 8.8: A single domain with magnetization close to +1 in a background where the magnetization is close to −1.

antikink are nearly independent. Under this assumption, the spatial dependence of the order parameter is, using Eq. (8.24),

$$m(x,t) \approx \tanh(x - x_1(t)) - \tanh(x - x_2(t)) - 1. \tag{8.57}$$

Let's now estimate how the interfaces move for this kink/antikink pair. Substituting the profile (8.57) into the TDGL equation (8.2) and keeping only the lowest-order terms, we find that asymptotically

$$\dot{x}_1 = -\dot{x}_2 \approx e^{-2(x_2 - x_1)}. \tag{8.58}$$

Consequently the domain length $L = x_2 - x_1$ shrinks according to $\dot{L} \approx -2e^{-2L}$ which gives $L(t) = \frac{1}{2}\ln\left[e^{2L(0)} - 4t\right]$, while the time for the domain to disappear is $\tau = \frac{1}{4}e^{2L(0)}$.

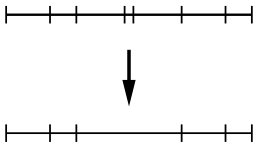


Figure 8.9: Extremal dynamics. The shortest domain merges with its two neighbors, while all other domains remain static.

This exponential length dependence of the shrinking time has profound consequences for the evolution of a heterogeneous domain array in one dimension. In the long-time limit, the shrinking time of the smallest domain is orders of magnitude shorter than that of the next smallest domain. As a result, the order in which domains merge with their neighbors becomes deterministic. This length ordering leads to the following simple extremal dynamics for TDGL domain evolution: (i) pick the smallest domain and merge it with its two nearest neighbors while keeping all other domains fixed; (ii) repeat *ad infinitum* (Fig. 8.9).

The domain length distribution

We now study how the domain length distribution evolves by this extremal dynamics. For simplicity, we assume that the total initial magnetization vanishes, so that the average length of domains of positive and negative magnetization are equal. The ensuing extremal dynamics coarsening resembles the Swendsen-Wang and Wolff cluster dynamics that were discussed in Sec. 7.5. However, in extremal dynamics, the natural time-like variable is the length ℓ of the shortest domain. Our goal is to determine $c(x,\ell)$, the density of domains of length x when the shortest domain has length ℓ. We choose the initial domain lengths from a continuous distribution so that each domain has a distinct length. The total density of domains is $\rho = \int_\ell c(x,\ell)\,dx$ and by definition $\int_\ell^\infty x\,c(x,\ell)\,dx = 1$.

Let's first show how the shortest domain plays the role of a time-like variable. When the shortest domain has length ℓ, suppose that a certain number of mergings occur so that all domains with lengths in the range $(\ell, \ell + \Delta\ell)$ merge with their neighbors so that the length of the shortest domain increases from ℓ to $\ell + \Delta\ell$. The density of the domains in this length range is $c(\ell,\ell)\Delta\ell$. Since there is a net loss of two domains in each merger, $\rho(\ell + \Delta\ell) = \rho(\ell) - 2c(\ell,\ell)\Delta\ell$. Thus the domain density obeys

$$\frac{d\rho}{d\ell} = -2c(\ell), \tag{8.59}$$

and in this rate equation the minimal length ℓ plays the role of a time.

The master equation for the domain length distribution may now be written in terms of this time-like variable. When the length of the shortest domain increases from ℓ to $\ell + \Delta\ell$, the length distribution $c(x,\ell)$ changes as follows:

$$c(x,\ell+\Delta\ell) - c(x,\ell) = \left[-2\,\frac{c(x,\ell)}{\rho} + \Theta(x-3\ell)\int_\ell^{x-2\ell} \frac{c(y,\ell)}{\rho}\,\frac{c(x-\ell-y,\ell)}{\rho}\,dy\right] c(\ell,\ell)\Delta\ell. \tag{8.60}$$

The first term on the right accounts for the loss of a domain of length x due to its merging. the factor of 2 stems from the fact that a domain of length x can be either to the left or to the right of the minimal domain. The second term accounts for the gain of a domain of length x due to the merging of three domains of lengths ℓ, y, and $x-\ell-y$. The Heaviside step function $\Theta(x-3\ell)$ enforces the condition that when the smallest domain has length ℓ, the smallest domain that can be created by merging must have length equal to at least 3ℓ. The last factor $c(\ell,\ell)\Delta\ell$ counts the number of merging events that occur as the minimal size increases from ℓ to $\ell+\Delta\ell$.

In close analogy with the discussion of cluster dynamics in Sec. 7.5, a remarkable feature of extremal evolution is that if the domains are initially uncorrelated, they remain uncorrelated at all times. Merging of domains do not affect their neighbors, nor are domains affected by their neighbors. Therefore the domain length distribution evolves according to the exact equation

$$\frac{dc(x,\ell)}{d\ell} = c(\ell,\ell)\left[-2\,\frac{c(x,\ell)}{\rho} + \Theta(x-3\ell)\int_\ell^{x-2\ell} \frac{c(y,\ell)}{\rho}\,\frac{c(x-\ell-y,\ell)}{\rho}\,dy\right] \tag{8.61}$$

for $x > \ell$. At this point, it is again convenient to introduce the normalized length density $P(x,\ell) = c(x,\ell)/\rho$ whose governing equation includes only the gain term:

$$\frac{dP(x,\ell)}{d\ell} = \Theta(x-3\ell)P(\ell,\ell)\int_\ell^{x-2\ell} P(y,\ell)P(x-\ell-y,\ell)\,dy. \tag{8.62}$$

Let's now investigate the asymptotic behavior of the domain length distribution. The ever-growing minimal domain length ℓ defines a basic scale, and we postulate that ℓ is the only length scale in the long time limit. Thus we assume that $P(x,\ell)$ approaches the scaling form

$$P(x,\ell) \simeq \ell^{-1}\Phi\left(x\ell^{-1}\right) \tag{8.63}$$

as $\ell\to\infty$, where the prefactor is fixed by the normalization condition $\int_1^\infty P(x,\ell)\,dx = 1$. As a result, the scaling function must satisfy the normalization condition $\int_1^\infty \Phi(z)\,dz = 1$. Substituting the scaling ansatz for $P(x,\ell)$ into the master equation (8.62), the scaling function obeys the nonlinear integro-differential equation

$$z\frac{d\Phi(z)}{dz} + \Phi(z) + \Theta(z-3)\Phi(1)\int_1^{z-2}\Phi(y)\Phi(z-1-y)\,dy = 0. \tag{8.64}$$

Given that the master equation involves a convolution, we again study the master equation in the Laplace domain. Here we introduce the Laplace transform $\phi(s) = \int_1^\infty \Phi(z)e^{-sz}dz$ with the lower limit of 1 because the convolution has this same lower limit. Multiplying (8.64) by e^{-sz} and integrating by parts, the Laplace transform then obeys the ordinary differential equation

$$s\frac{d\phi}{ds} = -\Phi(1)\left(1-\phi^2\right)e^{-s}, \tag{8.65}$$

with the boundary condition $\phi(0) = 1$. Expanding $\phi(s) = 1 + s\phi'(0)$ on the right hand side and evaluating the equality at $s = 0$ yields the normalized density of the shortest domains $\Phi(1) = 1/2$. Consequently, the asymptotic density of the shortest domains is given by

$$P(\ell) \simeq (2\ell)^{-1}. \tag{8.66}$$

Substituting $\Phi(1) = \frac{1}{2}$ into (8.65) and solving this equation yields the Laplace transform

$$\phi(s) = \tanh\left[\mathrm{Ei}\,(s)/2\right], \tag{8.67}$$

where $\mathrm{Ei}(x)$ is the exponential integral $\mathrm{Ei}(x) \equiv \int_x^\infty \frac{1}{u}\, e^{-u}\, du$.

The average domain length, $\langle x\rangle \simeq \langle z\rangle \ell$, follows from the small-argument behavior $\phi(s) \approx 1 - s\langle z\rangle$ and the asymptotic properties of the exponential integral[1] to give

$$\langle x\rangle \simeq 2e^\gamma\, \ell. \tag{8.68}$$

The ratio between the average length and minimal domain length approaches $2e^\gamma = 3.562144$. The conservation law $\rho\langle x\rangle = 1$ then gives the total domain density $\rho \simeq \frac{1}{2}e^{-2\gamma}\ell^{-1}$. Indeed, the density satisfies (8.59) $\frac{d\rho}{d\ell} = -2P(\ell)\rho \simeq \rho/\ell$.

Extremal properties of the domain distribution can be evaluated directly from the integro-differential equation (8.64). In the length range $1 < z < 3$, the integral drops out. The scaling function obeys the differential equation $z\frac{d}{dz}\Phi(z) = -\Phi(z)$ with the boundary condition $\Phi(1) = 1/2$, so

$$\Phi(z) = (2z)^{-1}, \qquad 1 \le z \le 3. \tag{8.69}$$

Interestingly, the normalized domain density becomes time-independent, $P(x,\ell) \to (2x)^{-1}$, in the (time-dependent) range $\ell < x < 3\ell$.

add exponential tail of distribution

The scaling analysis employed effectively resets the length of shortest domain to one after each merger. The similarity solution can be viewed as a *fixed-point* of this renormalization procedure. This technique is sometime termed *real space renormalization* or *strong disorder renormalization*.

Section on persistence has been temporarily commented out.

Also the subsection on the autocorrelation function has been commented out.

8.6 Nucleation and Growth

Nucleation-and-growth processes[2] are useful for modeling a remarkable wealth of phenomena. In this process, size-less islands (seeds) nucleate with a spatially-homogeneous rate $\gamma(t)$. The shape of the islands is spherical and their radius is $r(s)$ where s is the island lifetime. A point in space may be covered by multiple islands (figure 8.10).

Eventually, space becomes completely filled. The volume fraction, $\rho(t)$, the fraction of space covered by islands, can be calculated exactly. Its complement, $1-\rho(t)$, the fraction of uncovered space, is the probability that a given point in space, say the origin, remains uncovered at time t. For the origin to remain uncovered, no nucleation events may occur at a distance smaller than $r(t-\tau)$ from the origin at all times $\tau < t$; the corresponding volume is $V(t-\tau) = V_d[(r(t-\tau)]^d$ with $V_d = \pi^{d/2}/\Gamma(1+d/2)$ the volume of the d-dimensional unit hypersphere. Since the nucleation process is random, the probability that no nucleation events occur in all of these volumes is

$$1 - \rho(t) = \exp\left[-\int_0^t d\tau\gamma(\tau)V(t-\tau)\right]. \tag{8.70}$$

There are two important special cases: (i) Homogeneous nucleation where the nucleation rate is constant, $\gamma(t) = \gamma$; (ii) Instantaneous nucleation where all nucleation seeds appear at the same time, $\gamma(t) = \sigma\delta(t)$. For a constant growth velocity, $dr/dt = v$, the uncovered fractions are

$$1 - \rho(t) = \begin{cases} \exp[-\gamma U_d v^d t^{d+1}] & \text{homogeneous nucleation;} \\ \exp[-\sigma V_d (vt)^d] & \text{instantaneous nucleation.} \end{cases} \tag{8.71}$$

Here, $U_d = V_d/(d+1)$. The argument of the exponential equals the nucleation rate times the "excluded" volume, i.e., the volume of the space-time $(\mathbf{y},\tau)$ region where nucleation events affect the origin $(\mathbf{0},t)$. For homogeneous nucleation, this region is a d-dimensional hyper-cone of radius $v_0 t$ and height t; namely, a

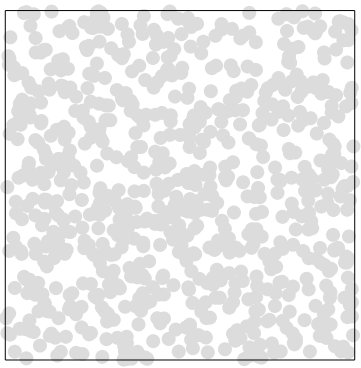


Figure 8.10: Nucleation-and-growth in two-dimensions (instantaneous nucleation).

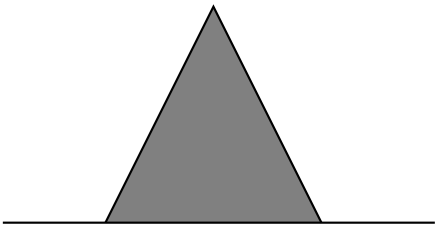


Figure 8.11: Homogeneous nucleation-and-growth in one-dimension. For a point to remain uncovered, no nucleation events may occur in a spacetime triangle.

triangle in one-dimension (Fig. 8.11), a cone in two-dimensions, etc. For instantaneous nucleation, the exclusion region is a d-dimensional sphere of radius $v_0 t$.

Multiple-point correlation functions can be calculated as well. To distinguish the covered phase from the uncovered one, we introduce the phase parameter

$$\phi(\mathbf{x},t) = \begin{cases} 1 & \text{if uncovered} \\ 0 & \text{if covered.} \end{cases} \tag{8.72}$$

The multiple-point correlation function $G_n(\mathbf{x}_1,\ldots,\mathbf{x}_n;t) = \langle \phi(\mathbf{x}_1,t)\cdots\phi(\mathbf{x}_n,t)\rangle$ equals the probability that all n points $\mathbf{x}_i$ remain uncovered at time t. This probability is evaluated following the very same reasoning used to evaluate the uncovered fraction. One should ensure that all nucleation events that can lead to covering of any of these points do not occur. The multiple point correlation function is

$$G_n(\mathbf{x}_1,\ldots,\mathbf{x}_n;t) = \exp\left\{ -\int d\mathbf{y} \int_0^t d\tau\, \gamma(\tau) \left[1 - \prod_{i=1}^{n} \Theta\big(|\mathbf{x_i}-\mathbf{y}| - r(t-\tau)\big) \right] \right\}, \tag{8.73}$$

[1] For small-x, $\mathrm{Ei}\,(x) \simeq -\gamma - \ln x$ with $\gamma = 0.577215$ the Euler constant.

[2] This process is refereed to as the Kolmogorov-Avrami-Johnson-Mehl or KAJM process.

where $\Theta(x)$ is the Heaviside step function; $\Theta(x) = 1$ for $x > 0$ and $\Theta(x) = 0$ for $x < 0$. The integration is over all possible nucleation sites $\mathbf{y}$ and nucleation times τ with the appropriate measure, the nucleation rate $\gamma(\tau)$. The integrand is zero when none of the points $\mathbf{x}_i$ are affected by a nucleation event at $(\mathbf{y}, \tau)$. It equals one when at least one of the points $\mathbf{x}_i$ would become covered due to nucleation at $(\mathbf{y}, \tau)$. We check that the one-point function $G_1(t) = G_1(\mathbf{x}, t) = 1 - \rho$ agrees with (8.70): $G_1(t) \equiv \exp[-\int_0^t d\tau\, \gamma(\tau) \int_{\mathbf{y}<r(t-\tau)} d\mathbf{y}]$; the spatial integral yields $V(t-\tau)$. The crucial feature that allows us to obtain the correlators is that nucleation events are uncorrelated, neither in space, nor in time.

In what follows, we consider constant growth velocities $dr/dt = v_0$. When all pairs of points are separated by distances larger than $2v_0t$, the integral (8.73) separates into n non-overlapping integrals, and the correlation function factorizes

$$G_n(\mathbf{x}_1, \ldots, \mathbf{x}_n; t) = G_1(\mathbf{x}_1, t) \cdots G_1(\mathbf{x}_n, t), \tag{8.74}$$

when $|\mathbf{x}_i - \mathbf{x}_j| > 2v_0t$ for all $i \neq j$. This simply reflects that there are no nucleation events that affect more than one point. Consequently, correlations vanish at large enough distances.

On the other hand, when there is a pair that is sufficiently close, correlations become nontrivial. Consider the pair correlation function $G_2(\mathbf{x}_1, \mathbf{x}_2)$ for homogeneous nucleation. The integral in (8.73) is proportional to the volume enclosed by the two hyper-cones originating at $\mathbf{x}_1$ and $\mathbf{x}_2$ (Fig. 8.12). The pair correlation function $G_2(\mathbf{x}, t) \equiv G_2(\mathbf{0}, \mathbf{x})$ may be written using the one point correlation $G_1(t) \equiv G_1(\mathbf{x}, t)$ in a way that manifests whether or not correlations are nontrivial

$$G_2(\mathbf{x}, t) = [G_1(t)]^2 \exp\left[\gamma U_d v^d t^{d+1} F_d\left(\frac{|\mathbf{x}|}{2v_0t}\right)\right]. \tag{8.75}$$

The characteristic function $0 \leq F_d(z) \leq 1$ quantifies the degree of correlation in d-dimensions. The characteristic function is normalized, $F_d(0) = 1$, and it vanishes, $F_d(z) = 0$, for $z \geq 1$.

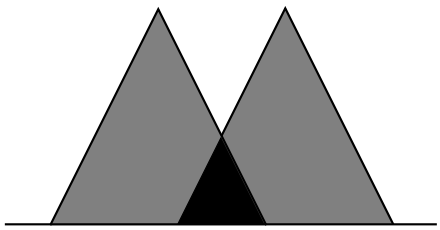


Figure 8.12: Pair correlation for homogeneous nucleation-and-growth in one-dimension. The overlap region between two spacetime triangle is highlighted.

The argument of the exponential in (8.73) equals the space-time volume of the overlap region involving two d-dimensional hyper-cones. For example, when $d = 1$, the overlap region is formed by two identical triangles of width $2v_0t$ and height t whose vertices are separated by distance $x < 2v_0t$. The overlap region is itself a triangle (figure 8.12) with width $2v_0t - x$ and height $t - x/(2v_0)$ and thus, its area is $(2v_0t - x)^2/(4v_0)$. Hence, the characteristic function is $F_1(z) = (1 - z)^2$ for $z < 1$.

Generally, the characteristic function $F_d(z)U_d = V_{\rm overlap}(z)$ is obtained from the overlap volume $V_{\rm overlap}(z)$ formed by two d-dimensional hyper-cones of radius 1 and height 1 whose vertices are separated by distance $2z$. Specifically,

$$F_d(z) = \begin{cases} 1 - z^2 & d = 1, \\ \frac{2}{\pi}\left[\cos^{-1} z - 2z\sqrt{1-z^2} + z^2 \ln\frac{1+\sqrt{1-z^2}}{z}\right] & d = 2, \\ \frac{1}{2}(1-z)^3(1+z) & d = 3. \end{cases} \tag{8.76}$$

In the complementary case of instantaneous nucleation, the pair correlation can be expressed as follows

$$G_2(\mathbf{x}, t) = [G_1(t)]^2 \exp\left[\sigma V_d v^d t^d F_d\left(\frac{|\mathbf{x}|}{2v_0t}\right)\right]. \tag{8.77}$$

Now, the characteristic function $F_d(z)V_d = V_{\text{overlap}}(z)$ is obtained from the overlap volume $V_{\text{overlap}}(z)$ formed by two d-dimensional hyper-cones of radius 1 separated by distance $2z$. For instance,

$$F_d(z) = \begin{cases} 1 - z & d = 1, \\ \frac{2}{\pi}\left[\cos^{-1} z - z\sqrt{1 - z^2}\right] & d = 2, \\ \frac{1}{2}(1 - z)^2(2 + z) & d = 3. \end{cases} \tag{8.78}$$

This geometric procedure extends to higher-order correlation functions. For example, the three-point correlation function is exponential in the total volume $V_{\text{total}} = V_1 + V_2 + V_3 - (V_{12} + V_{13} + V_{23}) + 2V_{123}$ where V_i is the single-point exclusion volume, V_{ij} is the volume of the exclusion region common to two points, and V_{ijk} is the volume of the exclusion region common to three points.

Chapter 9

REACTION KINETICS

In this chapter, we will discuss the time evolution of simple diffusion-limited reactions where one (or more) reactant species are converted into a product. There are two rates that control the overall reaction. The first is an intrinsic reactivity that specifies how quickly reactants in close proximity are converted to the product. The second is the rate at which the reactants actually meet; this transport mechanism is usually due to molecular diffusion. The interesting situation is the *diffusion-controlled* limit, in which conversion is quick and diffusion is the rate-limiting step. In contrast, in the *reaction-controlled* limit, reactants meet many times before a reaction actually occurs. We will also treat ballistic reactions where particles move at constant velocity between reactions.

The following prototypical reactions will be studied in this chapter:

- The catalytic reaction $A + C \to A + A + C$.
- Single-species annihilation $A + A \to 0$, where two diffusing particles annihilate when they meet.
- Coalescence, $A + A \to A$, where the reaction product is identical to the initial particles.
- Aggregation, $A_i + A_j \to A_{i+j}$. We will discuss this reaction in one dimension in contrast to the mean-field limit discussed in Chapter 4.
- Two-species annihilation, $A + B \to 0$, where different species annihilate when they meet.

These diffusion-controlled reactions have helped shaped the development of non-equilibrium statistical physics becuse of their simplicity and their phenomenological richness. As mentioned in Chapter 1, basic features of diffusion-controlled reactions can be obtained by scaling and dimensional analysis arguments. Our interest here is primarily in exact solution methods.

An important feature of such reactions is the crucial role played by the spatial dimension d. When d exceeds an upper critical dimension d_c, spatial fluctuations in the density of reactants are negligible and the reaction kinetics can be obtained by studying the mean-field limit. In spite of the irrelevance of spatial fluctions, the density distribution is not necessarily sharply peaked about its average value. Here the master equation for the density distribution provides considerable insights. When $d < d_c$, the spatial density of reactants becomes heterogeneous, leading to slower reaction kinetics compared to the mean-field limit. We will pay special attention to one-dimensional systems, where the master equation approach often yields exact solutions.

9.1 Catalytic Reaction $A + C \to A + A + C$

In the catalytic reaction $A + C \to A + A + C$, a new particle of type A is created whenever an A meets a catalyst C, while the catalyst remains unaffected by the reaction. We wish to understand the rate at which the number of A particles increases with time. In the mean-field limit, the average density of reactants $\langle n \rangle$ obeys the rate equation

$$\langle \dot{n} \rangle = k \langle n \rangle C, \tag{9.1}$$

where k denotes the reaction rate and C the catalyst density, and the angle brackets emphasize that the rate equation refers to the average number of particles. Without loss of generality, we take the product $kC = 1$. Then the solution to (9.1) is simply $\langle n(t)\rangle = \langle n(0)\rangle = e^t$. This result applies for a large and well-mixed system where both spatial fluctuations and fluctuations in the overall density are small. As we now show see, however, fluctuations in overall density are generally not small.

To account for realizations of the system in which the density is not close to the average density, we study the probability distribution $P_n(t)$ that the system contains n reactants at time t. For the catalytic reaction, the master equation for this distribution is just that of the Poisson process:

$$\dot{P}_n = (n-1)P_{n-1} - nP_n\,. \tag{9.2}$$

The meaning of the two terms is physically immediate: with rate proportional to $(n-1)P_{n-1}$ the particle number changes from $n-1$ to n, leading to a gain in the probability that there are n particles in the system. The loss term has a similar explanation. To solve this equation, we define the generating function $\mathcal{P}(z,t) = \sum_{n=1}^{\infty} P_n(t) z^n$, multiply Eq. (9.2) by z^n, and sum over all $n \geq 1$. After some simple manipulations, the generating function obeys the partial differential equation

$$\frac{\partial \mathcal{P}}{\partial t} = z(z-1)\frac{\partial \mathcal{P}}{\partial z}. \tag{9.3}$$

This equation may be solved by introducing the variable y via $dy = \frac{dz}{z(z-1)}$ (so that $y = \ln(1-z^{-1})$), to recast (9.3) as

$$\frac{\partial \mathcal{P}}{\partial t} = \frac{\partial \mathcal{P}}{\partial y}$$

with solution $\mathcal{P}(z,t) = F(y+t)$, where the function F is to be determined by the initial conditions.

For concretness, suppose that a single particle is initially in the system, $P_n(t=0) = \delta_{n,1}$, corresponding to $\mathcal{P}(z,0) = z$. Using $\mathcal{P}(z,0) = F(y)$ and inverting the expression above for $y(z)$ to give $z = (1-e^y)^{-1}$, we find $\mathcal{P}(z,0) = (1-e^y)^{-1}$. For $t > 0$, we then find $\mathcal{P}(z,t) = (1-e^{y+t})^{-1}$. We now expand this solution in a power series in z to extract $P_n(t)$. For this purpose, it is helpful to use the fact that the solution to the rate equation (9.1) for the single-particle initial condition is $\langle n\rangle = e^t$. Using this average density, the series expansion of the generating function is

$$\begin{aligned}
\mathcal{P} = \frac{1}{1-e^{y+t}} &= \frac{1}{1-\left(1-\frac{1}{z}\right)\langle n\rangle} \\
&= \frac{z}{\langle n\rangle}\,\frac{1}{1-z\left(\frac{\langle n\rangle - 1}{\langle n\rangle}\right)} \\
&= \sum_{n=1}^{\infty} z^n \frac{(\langle n\rangle - 1)^{n-1}}{\langle n\rangle^n}\,.
\end{aligned} \tag{9.4}$$

Now we may simply read off the solution:

$$P_n(t) = \frac{1}{\langle n\rangle}\left(1 - \frac{1}{\langle n\rangle}\right)^{n-1}\,. \tag{9.5}$$

In the long-time limit, $\langle n\rangle \gg 1$, so that the above expression approaches the Poisson form

$$P_n(t) \to \frac{1}{\langle n\rangle}\; e^{-n/\langle n\rangle} = e^{-t}\, e^{-ne^{-t}}\,. \tag{9.6}$$

Contrary to naive expectation, the number distribution is not a sharply peaked function about $\langle n\rangle$ because fluctuations in n are of the order of $\langle n\rangle$ itself. For the single-particle initial condition, it is simple to verify from (9.6) that

$$\sigma^2 \equiv \langle n^2\rangle - \langle n\rangle^2 = \langle n\rangle^2 - \langle n\rangle\,.$$

Thus the relative fluctuation is

$$\frac{\sigma^2}{\langle n\rangle^2} = 1 - \frac{1}{\langle n\rangle}\,,$$

which approaches 1 in the long-time limit. A surprising manifestion of these large fluctuations is the fact that although the average number of particles grows exponentially with time, the most likely event is that the system contains just a single particle! Thus even in the mean-field limit, finite-number fluctuations can be quite dramatic.

9.2 Single-Species Reactions

We now turn to the more substantial examples of: (i) single species annihilation, $A + A \to 0$, and (ii) single species coalescence, $A + A \to 0$.

Irreversible reaction: dimension dependence

When the reaction (either annihilation or coalescence) is irreversible, the spatial dimension plays a central role in determining the reaction kinetics. To see why this is the case, let's consider the mean-field rate equation for the average concentration $c \equiv \langle n \rangle$. Since two particles need to meet for a reaction to occur, the change in the concentration should be proportional to the probability that two reactants are in close proximity. Under the assumption of spatial homogeneity, this meeting probability factorizes into a product of single-particle densities. When particles do meet, their reaction rate is given by $k \propto Da^{d-2}$ (see Sec. 2.5 and Eq. (2.53) in particular), where D is the diffusion coefficient and a is the particle radius. From these arguments, the rate equation is:

$$\frac{dc}{dt} = -k\, c^2. \tag{9.7}$$

There is a difference of a factor of 2 in the reaction rates of annihilation and coalescence, but this difference is immaterial in the following discussion.

For $d \geq 2$, the reaction rate is an increasing function of particle radius and diffusivity. What happens for $d \leq 2$? There seems to be a problem because the reaction rate $k \propto Da^{d-2}$ *decreases* as the particle radius increases. As discussed in Sec. 2.5, the reaction rate now becomes

$$k(t) \propto \begin{cases} D^{d/2}\, t^{(d-2)/2} & d < 2, \\ 4\pi D/\ln t & d = 2, \\ Da^{d-2} & d > 2. \end{cases} \tag{9.8}$$

This dependence accounts for the fact that a random walk is certain to eventually return to its starting point so that the particle density near an absorbing point decreases with time. Using this reaction rate in the rate equation (9.7) then gives the asymptotic decay of the concentration for all spatial dimensions:

$$c(t) \propto \begin{cases} t^{-d/2} & d < 2, \\ t^{-1} \ln t & d = 2, \\ t^{-1} & d > 2. \end{cases} \tag{9.9}$$

The main point is that the density decays more slowly in $d \leq 2$ than the mean-field theory prediction. This slow decay is a manifestation of the depletion zone around each reactant.

Mean-field limit

While the rate equation (9.7) predicts that the average density asymptotically decays as t^{-1}, we can also study the *probability distribution* of the number of reactants to determine the relative importance of density fluctuations. As in the previous section, we let $P_n(t)$ be the probability that the system contains n particles. For irreversible annihilation, this probability distribution obeys the master equation

$$\dot{P}_n = \frac{1}{2}[(n+1)(n+2)P_{n+2} - n(n-1)P_n]\,. \tag{9.10}$$

Here we have set the reaction rate to 1. The first term accounts for the gain in $P_n(t)$ due to annihilation events in which the number of particles changes from $n+2$ to n. The rate of these events is proportional

to the number of AA pairs, namely $\frac{1}{2}(n+1)(n+2)$. The second term accounts for the complementary loss process in which $n \to n-2$. To solve this master equation, we again multiply by z^n and sum over all $n \geq 1$. After some simple steps, the generating function $\mathcal{P}(z,t) = \sum_{n=1}^{\infty} P_n(t) z^n$ obeys

$$\frac{\partial \mathcal{P}}{\partial t} = \frac{1}{2}(1-z^2)\frac{\partial^2 \mathcal{P}}{\partial z^2}. \tag{9.11}$$

To solve this equation we use separation of variables. We define $\mathcal{P}(z,t) = \mathcal{Z}(z)\mathcal{T}(t)$, substitute into Eq. (9.11), divide through by $\mathcal{P}$, and find

$$\frac{\dot{\mathcal{T}}}{\mathcal{T}} = \frac{1-z^2}{2}\frac{\mathcal{Z}''}{\mathcal{Z}} \equiv -\frac{1}{2}n(n-1)\,, \tag{9.12}$$

where the overdot and prime denotes differentiation with respect to time and z, resepctively. The solution for the time dependence is just exponential decay, while the z-equation

$$\mathcal{Z}'' + \frac{n(n-1)}{1-z^2}\mathcal{Z} = 0$$

is solved by the Gegenbauer polynomials of index $-\frac{1}{2}$, $C_n^{-\frac{1}{2}}(z)$; the appearance of this polynomial dictated the choice of the separation constant in (9.12). The general solution to (9.11) is then a linear combination of these elemental solutions

$$\mathcal{P}(z,t) = \sum_{n=0}^{\infty} A_n C_n^{-\frac{1}{2}}(z)\, e^{-n(n-1)t/2}\,,$$

where the coefficients A_n are determined by the initial conditions.

One dimension

One important realization of this reaction is the dynamics of the interfaces between domains of aligned spins in the one-dimensional Ising-Glauber model, as discussed in Sec. 7.2.

Perhaps the most elementary reaction is single-species annihilation, $A + A \to O$. It is convenient to consider this reaction when the reactants live on the sites of a regular d-dimensional lattice. Initially, the reactant density is c_0 and we allow each site to be occupied by at most one particle. Particles hop to a nearest-neighbor site with a constant rate, set to 1/2. When the destination site is occupied, annihilation occurs. Annihilation occurs with rate 1, twice the rate of hopping, because either of the two neighboring particles may hop. These hopping and annihilation events are illustrated in Figure 9.1 for one dimension.

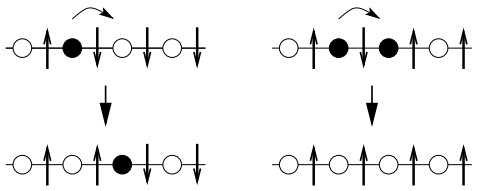


Figure 9.1: Single-species annihilation on a one-dimensional lattice: (left) hopping to an empty site with rate 1/2 and (right) annihilation with rate 1. Also shown is an equivalent representation in terms of Ising spins, where a pair of oppositely-oriented spins is equivalent to an intervening domain wall particle.

The reason why this problem in one dimension is so simple is that Glauber unknowingly already solved the problem! Let's recall our discussion in Sec. 7.2 of the $T = 0$ Ising model that is endowed with single spin-flip dynamics. Transitions that raise the energy are forbidden, while energy-lowering transition occur with rate 1, and energy-conserving transitions occur with rate 1/2. Identifying domain walls in the spin system with particles in the reaction, the two problems are identical (Fig. 9.1). Formally, the occupation number $n_i = 1$ or 0, that indicates whether a site is occupied or empty, is obtained from the corresponding spin configuration on the dual lattice via the transformation $n_i = (1 - s_i s_{i+1})/2$.

For simplicity, let's consider the completely occupied initial condition, $c(0) = 1$, corresponding to the antiferromagnetic initial condition in the equivalent spin system. From (7.38), the particle concentration c is (after identifying the particle concentration with the domain wall density $\equiv \rho$)

$$c(t) = I_0(2t)e^{-2t}. \tag{9.13}$$

From the asymptotics of the Bessel function, the concentration decays as

$$c(t) \simeq (4\pi t)^{-1/2}, \tag{9.14}$$

as $t \to \infty$. An intuitive way to obtain this result is to note that an isolated random-walk particle typically visits a region of size $x \sim t^{1/2}$ after a time t. when annihilation occurs, the number of particles that can remain within this length scale must be of the order of 1; if there were more particles in this region, they would have annihilated previously. Thus the typical spacing between particles is of the order of this diffusive length scale, and the concentration is the inverse of this scale. An important feature of this argument and also of the exact result is that the asymptotic density does not depend on the initial density.

Closely related to the particle density is the distribution of voids between particles. Using the nomenclature of chapter 6, a void of length n consists of n successive empty sites that is terminated at both ends by an occupied site. Let V_n be the distribution of voids of size n between two successive particles. Since the density decays as $t^{-1/2}$, the typical void length should grow as $t^{1/2}$. We therefore expect that the distribution of void lengths will depend only on the ratio of the length of a void to the typical void length. Consequently, the void length distribution should have the self-similar form

$$V_n(t) \simeq t^{-1}\Phi(nt^{-1/2}). \tag{9.15}$$

The time-dependent prefactor follows from the condition $\sum V_n \propto t^{-1/2}$. We can then use physical reasoning to find the asymptotic behavior of $\Phi(z)$. The density of minimal-size voids, those of length 0, is related to the density decay by $V_0 = -dc/dt \sim t^{-3/2}$. This asymptotic behavior is consistent with the above scaling form for V_n when the scaling function has the asymptotic behavior $\Phi(z) \sim z$ for $z \to 0$. This behavior is precisely what we found in the analysis of the void-size distribution for the zero-temperature Ising-Glauber model in Sec. 7.2. In addition to the linear vanishing of small-length voids, we also found that the void density decays exponentially at large distances

$$\Phi(z) \sim \begin{cases} z & z \ll 1, \\ e^{-z/z_*} & z \gg 1. \end{cases} \tag{9.16}$$

Steady state reaction

It is natural to study the influence of a steady input of reactants on annihilation. The input balances the loss of particles by annihilation so that a steady state is achieved. As we shall discuss, the precise nature of the input has a crucial effect on the long-time behavior. Consequently, the long-time state may actually represent thermodynamic equilibrium or the steady state retains a non-equilibrium character.

Let us first consider the simpler case where a pair of particles are inserted into neighboring sites of the system at a fixed rate h, while the reaction $A + A \to 0$ between nearest-neighbor pairs always occurs at rate 1. Why is adding pairs of particles at a fixed rate than merely adding single particles? To answer this question it is useful to view the annihilation reaction as creating an inert immobile product, that is, $A + A \to I$. The input of a pair is equivalent to the inert product splitting up into the pair of original A particles. Therefore pair input is nearly equivalent to the *reversible* annihilation reaction

$$A + A \leftrightarrow I.$$

There is a small caveat to this equivalence. In the case of pair input, once a pair of particles is created, they may diffuse away from their point of origin, after which an additional pair may then be created at this same origination point. In the reversible reaction, however, there cannot be multiple production of pairs from the vacuum, but only the production of a single pair from the breakup of a product particle. Thus the equivalence to the reversible reaction becomes exact in the limit of small input rate.

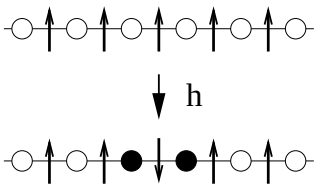


Figure 9.2: Equivalence between the input of a pair of particles at neighboring sites with a rate h and an energy raising event in the Ising-Glauber model.

In the nomenclature of the equivalent system of Ising spins with Glauber kinetics (see Fig. 9.2), the input of a particle pair corresponds to an energy-raising single spin-flip event. Consequently the input of pairs in the annihilation reaction is equivalent to the temperature being finite in the spin system.

Since the reaction is equivalent to the finite-temperature Ising-Glauber model, detailed balance is satisfied by construction. We now apply the detailed balance condition to derive the equilibrium density. Since $A + A \to I$ with rate 1 and $I \to A + A$ with rate h, the detailed balance condition is

$$1 \times P(\ldots 11 \ldots) = h \times P(\ldots 00 \ldots),$$

where $P(\ldots n_{i-1}, n_i, n_{i+1} \ldots)$ denotes the probability of the occupancy configuration $\{n_i\}$. Next we use the fact that for the Ising model in thermal equilibrium, the energy of each pair of spins is independent of all other spins. Therefore, the distribution of domain walls factorizes into a product over individual domain wall probabilities. Thus the detailed balance condition translates to the equation $c^2 = h(1-c)^2$, where c is the domain wall density. The steady-state density is therefore,

$$c(t) = \frac{h^{1/2}}{1 + h^{1/2}}. \tag{9.17}$$

One can verify that the same result also follows from (7.35) when $h = e^{-\beta}$. The steady-state density (9.17) is obvious *a posteriori* when one realizes that the sites are uncorrelated. The rate of particle gain equals the product of the probability of finding two vacant sites and the creation rate; similarly, the particle loss rate equals the probability of finding two occupied sites times the annihilation rate. These two processes lead to the rate equation $\frac{d\rho}{dt} = -\rho^2 + h(1-\rho)^2$ that again gives the steady-state density (9.17).

The steady state has a markedly different nature when particles are added one at a time

$$O \xrightarrow{h} A. \tag{9.18}$$

If a particle is added to an already occupied site, then annihilation is defined to occur immediately so that the outcome is an empty site. Thus the process $A \to 0$ also can occur, but its influence is negligible in the interesting limit of $h \to 0$. For this single-particle input, detailed balance cannot be satisfied as there are no processes that play the reverse role of annihilation and of input. Thus the system reaches a steady state with a fundamentally non-equilibrium character.

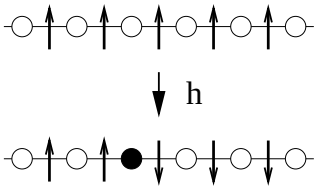


Figure 9.3: Equivalence between the input of a single with a rate h and a non-local energy raising event in the Ising-Glauber model in which all spins to the right of the domain wall flip.

We can determine the properties of this steady state by the Glauber formalism. When a single particle is created with rate h, the occupation n_i at the i^{th} site changes from 0 to 1. This single-particle creation corresponds to flipping all spins right of the i^{th} bond (Fig. 9.3), that is,

$$\dots s_{i-1}, s_i, s_{i+1}, s_{i+2} \dots \xrightarrow{h} \dots s_{i-1} s_i, -s_{i+1}, -s_{i+2} \dots . \tag{9.19}$$

Thus whenever a particle is added at the i^{th} with $1 \le i \le k$, all the spins s_j with $j > i$ also flip and the product $g_k = s_0 s_k$ changes sign. Therefore $g_k(t + \Delta t) = g_k(t)$ with probability $1 - (hk)\Delta t$ and $g_k(t + \Delta t) = -g_k(t)$ with probability $hk\Delta t$. The rate of change in the correlation function $G_k = \langle g_k \rangle$ due to the process (9.19) equals $-2hkG_k$. Adding this term to the master equation for the correlation function of the Ising-Glauber model Eq. (7.35), that also describes irreversible single-species annihilation, then gives

$$\frac{dG_k}{dt} = -2(1 + kh)G_k + G_{k-1} + G_{k+1} \tag{9.20}$$

for $k \ge 1$. The boundary condition is $G_0 = 1$.

In the steady state, the pair correlation function obeys

$$2(1 + kh)G_k(h) = G_{k-1}(h) + G_{k+1}(h). \tag{9.21}$$

Eq. (9.21) closely resembles the following recursion relation for the Bessel function of the first kind

$$\frac{2\nu}{x} J_\nu(x) = J_{\nu-1}(x) + J_{\nu+1}(x). \tag{9.22}$$

We can match this recursion with (9.21) be setting $\frac{2\nu}{x} = 2(1 + kh)$. This defines a one-parameter family of relations that connect (k, h) with (ν, x). A simple choice is $x = h^{-1}$ and $\nu = k + h^{-1}$, which then gives for the pair correlation function for $k \ge 0$,

$$G_k(h) = C\, J_{k+h^{-1}}(h^{-1}); \tag{9.23}$$

the prefactor $C = 1/J_{h^{-1}}(h^{-1})$ ensures the normalization $G_0 = 1$.

In the small-input limit, we make use of the asymptotic behavior of the Bessel function

$$J_\nu(\nu + x\nu^{1/3}) \sim (2/\nu)^{1/3} \mathrm{Ai}(-2^{1/3}x), \tag{9.24}$$

with $\mathrm{Ai}(x)$ the Airy function, to rewrite the particle density as

$$c = \frac{1}{2}(1 - G_1) \sim \frac{1}{2}\left[1 - \frac{\mathrm{Ai}((2h)^{1/3})}{\mathrm{Ai}(0)}\right]. \tag{9.25}$$

Expanding the Airy function to first order for small h, we obtain

$$c \sim 2^{-2/3} \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}\, h^{1/3} \approx 0.4593\, h^{1/3} \tag{9.26}$$

as $h \to 0$. For small h, the density is much larger compared to the $h^{1/2}$ dependence for the case of pair input (see Eq. (9.17)). The increased density for single-particle input arises because of the same spatial correlations between particles that occurs for irreversible annihilation.

The pair correlation function for single-particle input also differs substantially from the equilibrium correlation distribution. For large k, the recursion (9.21) for the correlation function reduces to

$$\frac{\partial^2 G(k)}{\partial k^2} = 2hk\, G(k). \tag{9.27}$$

For large k, this equation may be conveniently solved by using the WKB method and the result is (see highlight):

$$G_k \sim e^{-ak^{3/2}}, \tag{9.28}$$

with the constant $a = (8h/9)^{1/2}$. Thus correlations decay much more quickly with distance than the exponential decay (7.36) of the Ising-Glauber model at equilibrium.

The WKB method

The WKB method is a powerful analysis technique to obtain the asymptotic solution of a differential equation near an irregular singular point. A prominent such example is the equation (9.27), which can be written simply as $y'' = xy$. This equation also arises as the form of the time-independent Schrödinger equation near a classical turning point. At an irregular singular point, the dependence of y is faster than a power law and the standard approach to obtain the solution is to write it as the expansion $y = \exp[\phi_1(x) + \phi_2(x) + \ldots]$ and then solve for the expansion functions ϕ_n recursively. To leading order, we then obtain $(\phi_1')^2 = x$. There are two solutions to this equation, but the correct one is $\phi_1 = -\frac{2}{3}x^{3/2}$ which decays as $x \to \infty$. At the next level of approximation, we then obtain $\phi_2' = -\frac{1}{4}x^{-1}$. This yields the leading behavior

$$y \sim x^{-1/4} \exp\left[-\frac{2}{3}x^{3/2}\right]. \tag{9.29}$$

If one continues this method to the next level of approximation, one finds that all higher-order terms have the form of a vanishing correction to the leading behavior (9.29) as $x \to 0$.

As a counterpoint to the exact analysis given above, let's try to learn what we can about the steady state and the approach to the steady state by applying scaling. In the interesting small-input limit, we expect that the influence of the input will not be felt until the density decays to the point where the input represents a substantial perturbation. Thus for $h \to 0$, we expect that the density will decay as $t^{-1/2}$ in an intermediate-time regime that is large compared to the mean time for reactants to meet by diffusion, but small compared to the time between input events within a typical interparticle separation. However, at long times, the density should become constant. These two limiting behaviors may be encapsulated by the scaling ansatz

$$c(h,t) \sim h^\alpha \Phi(t\, h^\beta) \quad \text{with} \quad \alpha = \begin{cases} 1/3 & \text{pair input,} \\ 1/2 & \text{single-particle input.} \end{cases} \tag{9.30}$$

In this small-time limit, the input can be ignored and the density decays as $c \sim t^{-1/2}$. Assuming that $\Phi(z) \propto z^\gamma$ as $z \to 0$, we must have $\Phi(z) \sim z^{-1/2}$. Consequently, the exponent relation $\beta = 2\alpha$ must be satisfied to eliminate the dependence on the input rate. This reasoning gives

$$\beta = \begin{cases} 1 & \text{pair input,} \\ 2/3 & \text{single-particle input.} \end{cases} \tag{9.31}$$

Thus the relaxation for pair input is substantially slower that in single-particle input.

Finally, we can adapt the rate equation approach to give both the steady state and the time-dependent behavior under the influence of particle input. For pair input (equilibrium), we use the fact neighboring remain uncorrelated, $\frac{d}{dt}c = -c^2 + h$, while for the non-equilibrium steady-state, the effective reaction rate is proportional to the density and $\frac{d}{dt}c = -k(c)c^2 + h$. Using the effective reaction rates this generalizes Eq. (9.26) to arbitrary dimensions

$$c(h) \sim \begin{cases} h^{d/(d+2)} & d < 2, \\ h^{1/2}[\ln h^{-1}]^{1/2} & d = 2, \\ h^{1/2} & d > 2. \end{cases} \tag{9.32}$$

9.3 Coalescence $A + A \to A$

We now study the kinetics of the coalescence reaction $A + A \to A$ in one dimension. It is convenient to again to define the particle to live on the sites of a lattice and that each lattice site may be occupied by at most one particle. Particles hop to nearest neighbor sites with rate 1/2. If the destination site is occupied, the two particles coalesce, with the product having the same characteristics as the two initial particles. These processes can be represented by:

$$AO \xleftrightarrow{1/2} OA, \qquad AA \xrightarrow{1/2} OA \qquad AA \xrightarrow{1/2} AO. \tag{9.33}$$

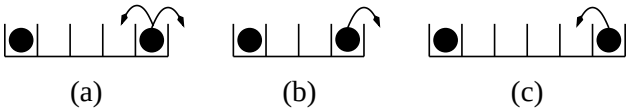


Figure 9.4: Changes in voids of size $n = 3$ due to hopping. Shown are the hopping events that lead to a loss (a) and the two types of gain processes ((b) & (c)).

A convenient way to analyze the coalescence reaction in one dimension is in terms of the voids between neighboring particles. Again, we use the terminology of chapter 6 that a void of length n is a string of n consecutive empty sites with the two sites at the end of the void occupied. The length of a void may either grow or shrink by 1 due to the hopping of a particle at the end of the void (Fig. 9.4). The dynamics of voids are "closed" because a void is affected only by the two particles at its boundary (all other particles are irrelevant!) This closure allows us to write a soluble equation for the void size distribution. Let V_n be the density of voids of size n. As illustrated in Fig. 9.4, the void size performs a random walk and the corresponding master equation for the void size density is (for $n > 0$)

$$\frac{dV_n}{dt} = -2V_n + V_{n+1} + V_{n-1}. \tag{9.34}$$

This equation can be extended to $n = 0$ by noting that the equation for the density of minimal size voids ($n = 0$), $\frac{dV_0}{dt} = -2V_0 + V_1$, can by put into the same form as (9.34) if one imposes the boundary condition $V_{-1} \equiv 0$.

For simplicity, consider the completely filled initial configuration, $V_n(0) = \delta_{n,0}$. As discussed in Sec. 7.2 (see especially Eq. (7.32)), the solution has the form $I_n(2t)e^{-2t}$, with $I_n(x)$ the modified Bessel function of the first kind. To satisfy the boundary condition $V_{-1}(t) = 0$ we use the image method. Thus we initially place a negative image charge of strength -1 at $n = -2$. Then the void density is

$$V_n(t) = [I_n(2t) - I_{n+2}(2t)]\, e^{-2t}. \tag{9.35}$$

We now use the fact that the particle density equals the total void density, $c = \sum_{k=0}^{\infty} V_k(t)$, since there is a one-to-one mapping between particles and voids. Consequently, the particle density is

$$c(t) = [I_0(2t) + I_1(2t)]\, e^{-2t}. \tag{9.36}$$

Asymptotically, the concentration decays algebraically, $c(t) \simeq (\pi t)^{-1/2}$, as in annihilation (see Eq. (9.14)). The amplitude of the asymptotic decay for coalescence is twice that of annihilation because one particle is lost in each coalescence reaction while two particles are lost in each annihilation.

Using the identity $I_{n-1}(x) - I_{n+1}(x) = \frac{2n}{x} I_n(x)$, we may simplify the result (9.35) for the void density to

$$V_n(t) = \frac{n+1}{t}\, I_{n+1}(2t)\, e^{-2t}. \tag{9.37}$$

In the long time limit, the void density becomes self-similar, following the same form (9.15) as in annihilation. This scaling form is consistent with both the typical void size $n \sim t^{1/2}$ and $c \sim t^{-1/2}$. The scaling function is

$$\Phi_{\rm coa}(z) = \sqrt{\frac{2}{\pi}}\, z\, e^{-z^2/2}. \tag{9.38}$$

At large distances, the void distribution has a Gaussian tail modified by an algebraic prefactor. Large voids are much less likely compared with annihilation, where there is an exponential decay. At small distances, the void density vanishes linearly, reflecting the fact that particles are effectively repelling each other. Particles enhance their survival rate by staying away from each other.

The void distribution shows: (i) correlations are generated dynamically and (ii) particle positions are correlated. Fluctuations are significant and must be taken into account because the likelihood of finding two neighboring particles is much smaller than c^2. The emergence of substantial fluctuations is responsible for the failure of the hydrodynamic approach.

Finite-Size Scaling

Often, we can understand long time asymptotics by determining the fate of a finite system. In coalescence, the final state of a finite system of size L is deterministic and consists of one particle. Consequently, the concentration must saturate at $c(L, t \to \infty) = L^{-1}$. We now assume that the time-dependent concentration is self-similar, *i.e.*, we postulate the scaling form

$$c(L,t) \simeq L^{-1}\Phi\left(\frac{L}{\sqrt{Dt}}\right).$$

The long time behavior dictates the scaling function behavior $\Phi(z) \sim 1$ as $z \to 0$. In the complementary limit $z \to \infty$, the behavior should be independent of system size and therefore $\Phi(z) \sim z$, thereby reproducing the asymptotic behavior $c(t) \sim (Dt)^{-1/2}$.

9.4 Aggregation $A_i + A_j \to A_{i+j}$

In aggregation, each site is either vacant or occupied by a cluster of mass i. Clusters hop to nearest-neighbor sites and when the target site is occupied, the aggregation event $A_i + A_j \to A_{i+j}$ occurs. Here A_k denotes a cluster of mass k. Let c_k be the density of clusters of mass k; mass conservation gives $\sum_k kc_k = 1$. We study spatially homogeneous situations with the monodisperse initial condition, $c_k(0) = \delta_{k,1}$. The basic question we want to answer is: what is the cluster mass distribution and how does it evolve in time? For the constant reaction kernel ($K_{ij} = 1$), aggregation is exactly soluble in one dimension by generalizing the empty interval method.

Irreversible reaction

To determine the cluster mass distribution, it is convenient to introduce auxiliary variables that quantify the amount of mass within an interval of a given size. We define Q_n^k as the probability that the total mass contained in n consecutive sites equals k. By construction $\sum_{k=0}^{\infty} Q_n^k = 1$, and the probability that there is no mass in the n interval, Q_n^0 is just the empty interval probability that was first introduced in our discussion of adsorption phenomena in chapter 6; thus $E_n \equiv Q_n^0$. The fundamental cluster mass distribution, namely the probability to have a cluster of mass k is simply the probability that there is a mass k in an interval of length 1; thus $c_k = Q_1^k$.

The feature that makes aggregation soluble is that the interval probabilities Q_n^k evolve according to the very same discrete diffusion equation that also governs the void and the empty interval densities! The derivation, though, is more delicate because of the need to track both the interval length and the mass contained within the interval. We now require the conditional probability $\tilde{Q}_n^k$ for n consecutive sites to contain a total mass k are followed by an empty site. To write the master equation for Q_n^k, we detail the changes in this quantity due to hopping events at the right boundary (9.5). A similar set of contributions arise at the left boundary.

Intervals of size n and mass k are gained (+) and lost (−) with the following rates:

+ The mass in an n-interval is smaller than k and a cluster hops into this interval to make the final mass equal to k. Thus the mass contained in the interval of length $n+1$ must equal k. The rate of this event equals $[Q_{n+1}^k - \tilde{Q}_{n+1}^k]/2$. The difference of the Q's accounts for the probability that a mass k is contained in an interval of length $n+1$ with the last site on the right of this interval being occupied. The factor $1/2$ accounts for this last cluster hopping to the left to create an interval of length n that contains mass k.

+ The interval mass is larger than k and a cluster at the end of the interval hops out so that the final mass equals k. The rate for this event is $[Q_{n-1}^k - \tilde{Q}_{n-1}^k]/2$.

− The interval mass equals k and a cluster hops into it with rate $-[Q_n^k - \tilde{Q}_{n+1}^k]/2$.

− The interval mass equals k and a cluster hops out of it with rate $-[Q_n^k - \tilde{Q}_{n-1}^k]/2$.

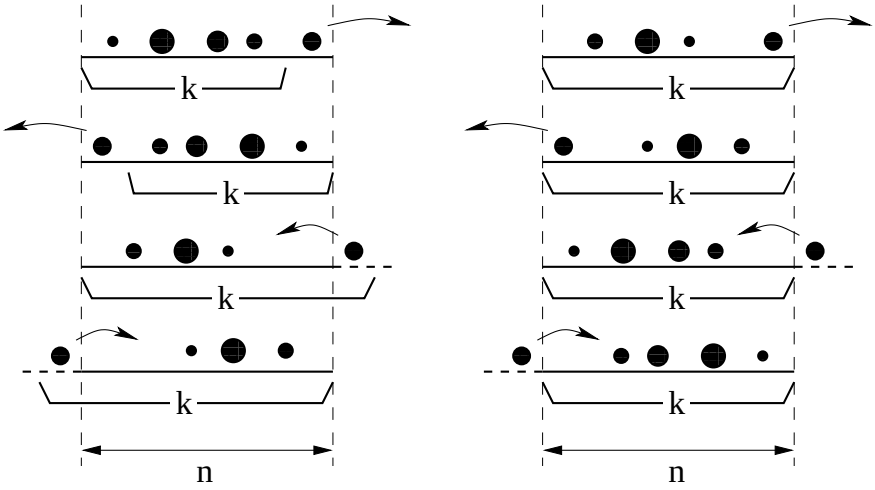


Figure 9.5: Configurations that contribute to the change in the set of states in which a total mass k is contained in an interval of length n. The left-hand side shows the four processes that lead to a gain in Q_n^k, and the right-hand side shows the four loss processes.

Adding all these transition rates, the conditional probabilities miraculously cancel! By including the identical contribution at the second boundary, the evolution of the empty interval density is again described by the discrete diffusion equation

$$\frac{dQ_n^k}{dt} = -2Q_n^k + Q_{n-1}^k + Q_{n+1}^k \tag{9.39}$$

for all $k \geq 0$. The boundary condition is $Q_0^k(t) = 0$ for $k > 0$ (indeed, $E_0 = Q_0^0 = 1$) and we choose the initial condition $Q_n^k(0) = \delta_{n,k}$, corresponding to each site of the lattice initially occupied by a monomer. These equations have two remarkable features: they are (i) closed, and (ii) uncoupled (different masses not coupled). This empty interval method can be generalized to both spatially-inhomogeneous situations and monomer input. The solution can now be obtained by using the image method and gives $Q_n^k(t) = [I_{n-k}(2t) + I_{n+k}(2t)]e^{-2t}$. From this result, the cluster mass density $c_k = Q_1^k$ is

$$c_k(t) = [I_{k-1}(2t) - I_{k+1}(2t)]e^{-2t}. \tag{9.40}$$

Incidentally, the cluster densities are identical to the void densities (9.35), $c_k = V_{k+1}$. Intuitively, one may imagine that the entire mass in a void is contained by the cluster to its right.

Asymptotically, the cluster size distribution is

$$c_k(t) \simeq \frac{k}{\sqrt{\pi t^3/2}}\, e^{-k^2/2t}. \tag{9.41}$$

This distribution can be written in the scaling form, $c_k \sim t^{-1}\Phi(kt^{-1/2})$ with the scaling function (9.38). The scaled mass distribution differs significantly from the corresponding result for constant-kernel aggregation in the mean-field limit, $\Phi(z) = e^{-z}$ (see Eq. (4.11)). This mean-field result holds above the critical dimension $d > 2$. In one dimension there is a depletion of smaller than typical clusters $k \ll t^{1/2}$ and the decay at large masses is sharper, too: $\Phi(z) \sim e^{-z^2/2}$. Finally, notice that by summing the cluster mass distribution over all masses, we reproduce the density for coalescence, $c_{\rm coa} = c_1 + c_2 + c_3 + \ldots$. Similarly, by summing over odd sizes only, we reproduce the density for the annihilation case, $c_{\rm ann} = c_1 + c_3 + c_5 + \ldots$.

Aggregation with input

What happens when we now add monomers to the system at a constant rate? In our discussion of chapter 4 about aggregation with input in the mean-field limit, we found that a non-trivial steady state was created.

For the case of a constant reaction kernel, the steady-state mass distribution, $c_k(t \to \infty)$, had a $k^{-3/2}$ tail over a mass range $1 \ll k \ll t^2$. Our goal is to determine the corresponding behavior of steady-state aggregation in one dimension.

We first generalize the lattice description of aggregation by adding monomers to the system at a constant rate at each site: $0 \to A$ with rate h. In spite of this extra ingredient in the dynamics, it is still possible to write and solve the equations for the empty interval probabilities. The effect of input on the empty interval probability is quite simple: if an n interval contains mass $k-1$ and an input event occurs, there is a gain in Q_n^k. The rate at which mass is added to this interval is just $h\,n$; this is proportional to the interval size because input may occur at any of the sites. Similarly, if the interval contains mass k, input causes the loss of Q_n^k. Thus the master equation for Q_n^k is

$$\frac{dQ_n^k}{dt} = -2Q_n^k + Q_{n-1}^k + Q_{n+1}^k + h\,n\left[Q_n^{k-1} - Q_n^k\right]. \tag{9.42}$$

This equation holds for all $k > 0$ with the boundary condition $Q_n^{-1}(t) = 0$. While the equations are easy to formulate, the full time-dependent behavior is harder to obtain than in irreversible aggregation because the interval probabilities for different contained masses and different lengths are now all coupled.

However, the situation is much simpler in the steady state. In this case, we introduce the generating function, $Q_n(z) = \sum_k Q_n^k e^{kz}$, to convert Eq. (9.42) to

$$Q_{n-1}(z) + Q_{n+1}(z) = [2 + hn(1 - e^z)]\,Q_n(z). \tag{9.43}$$

This recursion formula is the same as that for the Bessel function (Eq. (9.22)) when the index is properly matched. Thus following exactly the same line of reasoning as that leading to Eq. (9.23), The solution is now

$$Q_n(z) = \frac{J_{n+g^{-1}}(g^{-1})}{J_{g^{-1}}(g^{-1})}, \tag{9.44}$$

with $g \equiv g(z,h) = h(1-e^z)/2$. We are interested primarily in the large-k behavior of the mass distribution; this limit corresponds to the small-z behavior of the generating function. Thus we use the approximation $g \approx hz/2$ and the asymptotic formula (9.24) for the Bessel function to obtain

$$Q_1(z) \sim \frac{\mathrm{Ai}((2g)^{1/3})}{\mathrm{Ai}(0)} \sim 1 - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}(hz)^{1/3}. \tag{9.45}$$

This leading behavior $Q_1(z) = 1 - \text{const.} \times (hz)^{1/3}$ then implies an algebraic decay of the mass distribution

$$c_k \sim k^{-4/3}, \tag{9.46}$$

for $k \gg 1$. The exponent differs from the mean-field theory prediction (4.75), $c_k \sim k^{-3/2}$. Even though the mean-field theory fails quantitatively, it is still extremely valuable because it helps us articulate "what to expect". When there is input, mean-field theory predicted a power-law decay of the mass distribution for large masses, and this is what also occurs in one dimension. Similarly, for irreversible aggregation, mean-field theory predicts scaling behavior and a rapidly decaying tail of the mass distribution, again in qualitative accord with the behavior in one dimension.

Random River Networks

Aggregation with input is equivalent to the classic Scheidegger random river network model (Fig. 9.6). In this model, the downstream position along a river is equivalent to the time and the lateral meandering of a river is equivalent to a one-dimensional random walk. The source of a minimal size river is then equivalent to the injection of a unit mass. The meandering of a river is captured by a random walk and the merging of two rivers is equivalent to aggregation, as the flow rate of the combined rivers equals that of the two tributaries. The river network is therefore equivalent to the space-time diagram of aggregation with input. Let h be the river's depth. It's defining edges perform two independent random walks, so the river height distribution equals the first passage probability, $p(h) \sim h^{3/2}$. The river size k is proportional to the area of its drainage basin. The area scales as the depth times the height and since the width is diffusive ($\sim h^{1/2}$), the river size is $k \sim h \times h^{1/2} \sim h^{3/2}$. The size distribution

$$p(k) = p(s)\frac{ds}{dk} \sim k^{-4/3} \tag{9.47}$$

obtained from this heuristic argument therefore agrees with the asymptotic result (9.46).

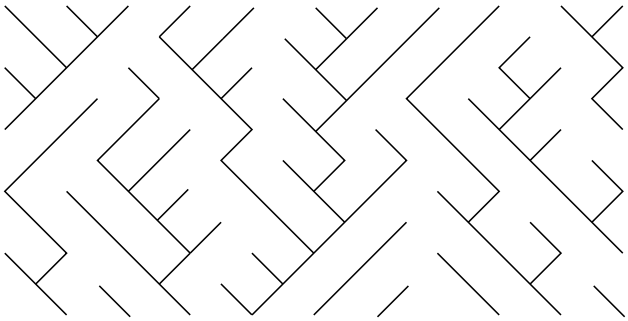


Figure 9.6: A river networks as aggregation with input. Shown are is the position versus time of the aggregates. This pictures depicts a "cellular automata" realization: a discrete time version where vacant site are immediately filled by particles. The lattice has been tilted by 45°.

9.5 Two Species Annihilation $A + B \to 0$

The reaction

$$A + B \to 0$$

is known as *two-species annihilation*. Here, same-species particles do not interact, while particles of the opposite species annihilate in pairs. Physical examples of this reaction include the annihilation of electron-hole pairs in a semiconductor or the annihilation of matter with antimatter in a cosmological setting. Perhaps the most striking aspect of diffusion-controlled two-species annihilation is that the density decays as $t^{-d/4}$ for spatial dimension $d < 4$ for equal initial densities of the two species. This decay is much slower than the rate equation prediction of a t^{-1} decay and also slower than the $t^{-d/2}$ decay of single-species reactions for $d < 2$. The basic feature that gives rise to this anomalously slow kinetics is that the reactants organize into a coarsening mosaic of single-species domains (fig. 9.7). As a result, reactions can occur only near domain boundaries, rather than uniformly throughout the system. The inherent heterogeneity of the reaction leads to slow kinetics.

While the $t^{-d/4}$ density decay has been proven by exact analysis methods, this approach involves mathematical techniques that lie outside the scope of this book. Thus in this chapter we will primarily discuss qualitative approaches to determine the many interesting physical properties of two-species annihilation. While these approaches lack mathematical rigor, they are intuitive and easy to appreciate.

Density decay

Let us first give a back-of-the-envelope argument for the long-time behavior of the concentration $c(t)$ in terms of local density fluctuations. Consider a finite volume of linear dimension L. The number of particles of each species in this volume is given by

$$N_{A,B} = c(0)L^d \pm \sqrt{c(0)}\; L^{d/2}. \tag{9.48}$$

Here the $\pm$ in the second term signifies that, in a finite volume L^d, $N_{A,B}$ has fluctuations that are of the order of $\sqrt{c(0)}\; L^{d/2}$, in which the amplitude of this term is of order 1 and the sign that fluctuates from realization to realization. We now focus on the symmetric system where the two species are initially present in equal numbers. Then the initial difference in the number of A and B particles in the volume is

$$N_A - N_B \approx \pm\sqrt{c(0)}\; L^{d/2}. \tag{9.49}$$

Again, the coefficient of the second term should be understood as a number that is of the order of one and whose sign is equally likely positive or negative.

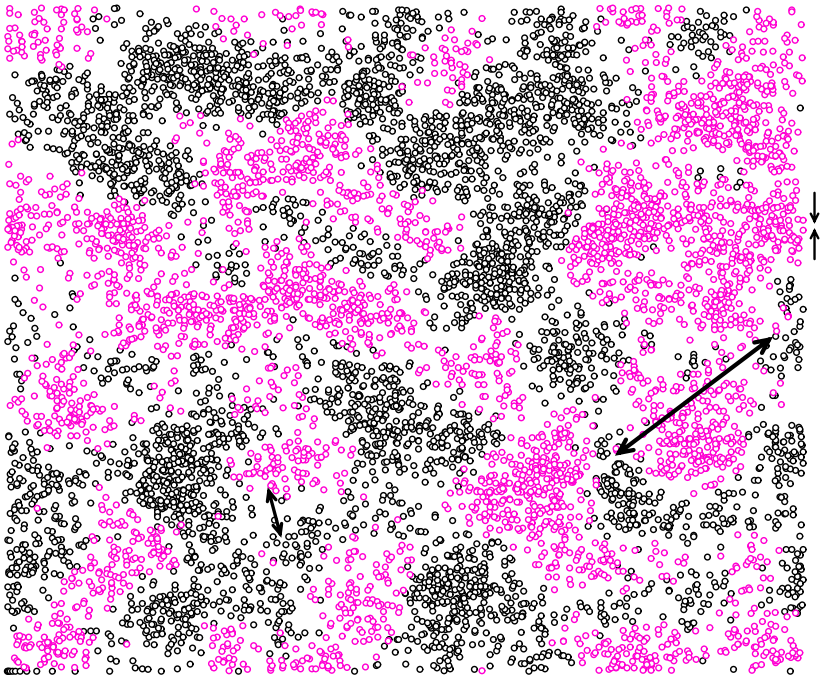


Figure 9.7: Snapshot of the particle positions in two-species annihilation in two dimensions, now showing the basic length scales of the system: the domain size, which scales as $t^{1/2}$, the interparticle spacing, which scales as $t^{1/4}$, and the depletion zone between domains, which scales as $t^{1/3}$ in two dimensions.

Roughly speaking, $N_A - N_B$ remains nearly constant during the time $t_L \sim L^2/D$ that it takes for a typical particle to traverse the volume by diffusion. After a time t_L, sufficient time has elapsed that particles have had time to annihilate with a member of the opposite species. Thus the "extensive" part of the particle number (the first term in Eq. (9.48)) will be eliminated, leaving behind the local majority species in the domain. By conservation of the difference of $N_A - N_B$, the number of particles in this local majority, $N_>(t_L)$, is of the order of is of order $\sqrt{c(0)}\, L^{d/2}$. Finally, by eliminating L in favor of t, we obtain

$$c(t) \approx N_>(t)/L^d \sim \sqrt{c(0)}\; (Dt)^{-d/4}, \qquad (d \leq 4). \tag{9.50}$$

Let us give a somewhat better grounded argument for anomalous $t^{-d/4}$ decay of the density by again focusing on the local density difference $\delta(\mathbf{r}, t) \equiv c_A(\mathbf{r}, t) - c_B(\mathbf{r}, t)$. The concentration of each species evolves

by the diffusion-reaction equation

$$\frac{\partial c_{A,B}(\mathbf{r},t)}{\partial t} = D\nabla^2 c_{A,B}(\mathbf{r},t) + R,$$

where R denotes the reaction term. We leave the reaction term unspecified because the density difference $\delta(\mathbf{r},t)$ evolves only by pure diffusion, $\frac{\partial \delta}{\partial t} = D\nabla^2\delta$. Consequently, the Fourier transform of the density difference is simply $\delta(\mathbf{k},t) = \delta(\mathbf{k},t=0)e^{-Dk^2t}$. At long times, there is minimal coexistence of A's and B's in the same spatial region because of the existence of domains. Thus $[c_A(\mathbf{r},t) - c_B(\mathbf{r},t)]^2 \approx 2c_A(\mathbf{r},t)^2$, so that

$$\begin{aligned}\int c_A(\mathbf{x},t)^2\, d\mathbf{x} &\approx \frac{1}{2}\int |\delta(\mathbf{k},0)|^2 e^{-Dk^2t}\, d\mathbf{k} \\ &\propto (Dt)^{-d/2}\int |\delta(\mathbf{q}/(Dt)^{1/2},t)|^2\, e^{-q^2}\, d\mathbf{q}. \end{aligned} \tag{9.51}$$

For a random initial condition, $|\delta(\mathbf{k},t=0)|^2 = N$ for all $\mathbf{k}$, since the mean-square involves the sum of N random unit vectors. Thus the integral over $\mathbf{q}$ in Eq. (9.51) is independent of t, so that $\langle c_A(\mathbf{x},t)^2\rangle \sim \frac{N}{V}\,(Dt)^{-d/2}$. Finally the assumption of no cross correlations implies that $\langle c_A(\mathbf{x},t)^2\rangle \cong \langle c_A(\mathbf{x},t)\rangle^2$ and back 1 is reproduced. Notice that the random initial condition is a crucial aspect for obtaining the anomalous slow decay. In particular, for correlated initial conditions with no long-wavelength fluctuations in $\delta(\mathbf{x},t)$, the integral over $\mathbf{q}$ will vanish as $t \to \infty$, thus invalidating the above reasoning.

We conclude that a homogeneous system that is equally populated by A and B particles evolves into a continuously growing mosaic of single-species domains. The identity of each domain is determined by the local majority species in this same spatial region in the initial state. At time t, these domains will be of typical linear dimension $\sqrt{Dt}$, within which only the species in the local majority remains, with concentration $\sqrt{c(0)}\;(Dt)^{-d/4}$.

The spontaneous formation of domains breaks down for $d > 4$, however, because single-species domains become transparent to an invader of the opposite species. Consider, for example, the fate of a single A particle that is placed at the center of a B domain of linear dimension L and local concentration therefore of order $L^{-d/2}$. The impurity needs L^2 time steps to exit the domain, during which L^2 distinct sites would have been visited (again assuming $d > 4$). At each site, the A particle will react with probability of the order of the B concentration, $L^{-d/2}$. Therefore the probability that an A particle reacts with any B particle before exiting this domain is of order $L^{(4-d)/2}$. Since this probability vanishes as $L \to \infty$ when $d > 4$, a domain is unstable to diffusive homogenization and the system as a whole therefore remains spatially homogeneous.

Spatial organization

The above arguments suggest that two lengths are needed to characterize the reactant distribution in one dimension: the linear dimension of a typical domain, $L \propto (Dt)^{1/2}$, and the typical interparticle spacing, which scales as $c(t)^{-1} \propto t^{1/4}$. Surprisingly, there is yet another fundamental length scale in the system—the typical distance between AB closest-neighbor pairs, ℓ_{AB}. The length ℓ_{AB} characterizes the gap that separates adjacent domains (Fig. 9.7). This gap is the fundamental control factor in the kinetics, since each reaction event involves diffusion of an AB pair across a gap.

To determine the evolution of ℓ_{AB}, let's first consider the simplest case of one dimension. We now reformulate the kinetics specifically in terms of the AB gap distance. Let c_{AB} denote the concentration of closest-neighbor AB pairs. Typical AB pairs react in a time $\Delta t \sim \ell_{AB}^2/D$. Since the number of reactions per unit length is of order c_{AB}, the rate of change of the overall concentration is

$$\frac{\Delta c}{\Delta t} \approx \frac{dc}{dt} \approx -\frac{c_{AB}}{\ell_{AB}^2/D}. \tag{9.52}$$

Now $\frac{dc}{dt}$ is known from $c(t)$ itself, while in one dimension, $c_{AB} \propto (Dt)^{-1/2}$, since there is one AB pair per domain of typical size $(Dt)^{1/2}$. Using these results and solving for ℓ_{AB} gives

$$\ell_{AB} \propto c(0)^{-1/4}\,(Dt)^{3/8}. \tag{9.53}$$

The fact that $\ell_{AB} \gg \ell_{AA}$ is a manifestation of the effective repulsion between opposite species. If one squints at Fig. 9.7, this inequality between ℓ_{AB} and ℓ_{AA} should be visually apparent.

The above results can be generalized to spatial dimension $1 \leq d \leq 2$. The time dependence of ℓ_{AB} still follows by applying (9.52), since it holds whenever random walks are compact (see the discussion in the two paragraphs following Eq. (2.39) for a definition of compact random walks). We now assume that the interface of a single-species domain remains relatively smooth, so that a domain of linear dimension ℓ will have an interface area of $t^{(d-1)/2}$. Assuming that the particles in this interfacial zone are separated by a distance of the order of ℓ_{AB}, irrespective of identity, it is straightforward to obtain

$$\ell_{AB} \propto t^{(d+2)/[4(d+1)]}, \qquad c_{AB}(t) \propto t^{-d(d+3)/[4(d+1)]}, \tag{9.54}$$

which gives $\ell_{AB} \sim t^{1/3}$ and $c_{AB}(t) \sim t^{-5/6}$ in $d = 2$. For $d > 2$, the transience of random walks implies that two opposite species particles within a region of linear dimension ℓ_{AB} will react in a time of order ℓ_{AB}^d (rather than ℓ_{AB}^2). Consequently, (9.52) should be replaced by

$$\frac{\Delta c}{\Delta t} \approx -\frac{c_{AB}}{\ell_{AB}^d}. \tag{9.55}$$

This relation, together with the assumption of a smooth interfacial region between domains, gives, in $d > 2$ dimensions

$$\ell_{AB} \approx t^{d+2/[4(2d-1)]}, \qquad c_{AB} \approx t^{-d^2+5d-4/[4(2d-1)]}. \tag{9.56}$$

These coincide with (9.54) at $d = 2$, but yield $c_{AB} \approx t^{-1}$ and $\ell_{AB} \approx t^{1/4}$ for $d = 3$. The latter represents the limiting behavior where ℓ_{AB} becomes of the same order as ℓ_{AA}. Thus the non-trivial scaling of interparticle distances disappears in three dimensions and above.

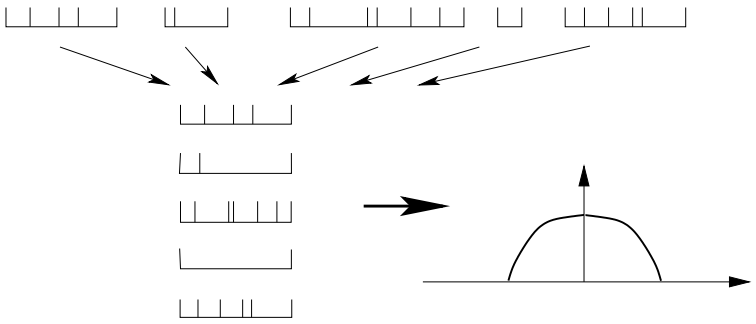


Figure 9.8: Construction of the microcanonical domain profile from the reactant positions (top line). Each domain is first scaled to a fixed length (lower left) and then their densities are superposed (lower right).

Much insight can be gained by studying the average density profile of a single domain. Consider the "microcanonical" density profile, $P^{(M)}(x)$, defined as the probability of finding a particle at a scaled distance x from the domain midpoint, when each domain is first scaled to a *fixed* size (Fig. 9.8). The resulting distribution is similar to the long-time probability distribution for pure diffusion in a fixed size absorbing domain. In contrast, for two-species annihilation, particles in a single domain are confined by absorbing boundaries which recede stochastically as $\sqrt{t}$ – the typical domain size. While the probability distribution inside such a stochastically evolving domain has not been solved, one can solve the related problem of a particle inside a deterministically growing domain $[-L(t), L(t)]$ with $L(t) \propto t^{1/2}$, The adiabatic approximation marginally applies in this case [16], and the density profile has the form $\cos(\pi x/L(t))$. This simple-minded modeling provides a useful framework to understand the domain profile in the reacting system.

Although determined by interactions between *opposite* species, this inhomogeneous domain profile governs the distribution of interparticle distances between *same* species. Particles are typically separated by a distance which grows as $t^{1/4}$ within the core of the domain, but systematically become sparser as the domain

interface is approached. The subregions of "core" and "interface" each comprise a finite fraction of the domain. These essential features of the profile may be accounted for by the trapezoidal form (Fig. 2(b)),

$$\rho(z) \equiv c(x,t)\, t^{1/4} = \begin{cases} \rho_0, & |z| \le z^*; \\ \rho_0(1-|z|), & z^* < |z| < 1-\epsilon. \end{cases} \tag{9.57}$$

Here $z \equiv x/L(t)$ is the scaled spatial co-ordinate, with $x \in [-L(t), L(t)]$, and ρ_0 and $z^* \lesssim 1$ are constants. The upper limit for $|z|$ on the second line of Eq. (9.57) reflects the fact that there are no particles within a scaled distance of $\epsilon \equiv \ell_{AB}/L(t) \sim t^{-1/8}$ from the domain edge. The linear decay of the concentration near the domain edge arises from the finite flux of reactants which leave the domain. Thus, the local nearest-neighbor distance is $\rho(z)^{-1}$, with $\rho(z) = \rho_0$ in the core ($|z| \le z^*$), but with $\rho(z) = \rho_0(1-|z|)$ near the boundary and the time dependence of the reduced moments of the AA distance distribution are

$$M_n \equiv \langle \ell_{AA}^n \rangle^{1/n} = \left(\int_0^\infty x^n\, P_{AA}(x,t)\, dx \right)^{1/n}, \tag{9.58}$$

$$\approx t^{1/4} \times \left(2\int_0^{z*} \frac{dz}{\rho_0^n} + 2\int_{z*}^{1-\epsilon} \frac{dz}{\rho_0^n (1-z)^n} \right)^{1/n}, \tag{9.59}$$

$$\sim \begin{cases} t^{1/4}, & n < 1; \\ t^{1/4} \ln t, & n = 1; \\ t^{(3n-1)/8n}, & n > 1. \end{cases} \tag{9.60}$$

For $n < 1$, the dominant contribution to M_n originates from the ρ_0^{-n} term in the parentheses, while for $n \ge 1$, the term involving $\rho_0^{-n}(1-z)^{-n}$ dominates, with the second term giving a logarithmic singularity at the upper limit for $n = 1$. Thus the large-scale modulation in the domain profile leads to moments $M_n(t)$ which are governed both by the gap length ℓ_{AB} and ℓ_{AA}. As $n \to \infty$, the reduced moment is dominated by the contribution from the sparsely populated region near the domain periphery where nearest-neighbor particles are separated by a distance of order $t^{3/8}$.

9.6 The Trapping Reaction $A + T \to T$

At first sight, the trapping reaction seems to be even simpler than annihilation or coalescence because trapping is essentially a single-particle problem. The system is populated by randomly-distributed static traps Fig. 9.9. Static traps are randomly distributed in space and independent particles freely diffuse in this medium. Whenever a diffusing particle hits a trap it is immediately and permanently trapped. What is the probability $S(t)$ that a particle "survives" until time t?. At the most naive level, one might argue that one can replace any realization of the trapping medium by an effective average medium with a constant trapping rate. This would suggest that the density of survivors should decay exponentially with time. Surprisingly, this naive expectation is wrong and for interesting reasons. As we shall discuss, extreme fluctuations in the spatial distribution of traps in the form of large trap-free regions give rise to a slower decay of the survival probability. However, this anomalously slow decay manifests itself only when the density has decayed to an astronomically small value. One has to be careful to understand what may be of fundamental theoretical interest and what may be experimentally relevant.

Exact solution in one dimension

The essence of the problem can be appreciated already in one dimension where we can obtain the exact solution. A diffusing particle "sees" only the absorbing interval defined by the nearest surrounding traps. We can therefore adapt the solution for the concentration inside an absorbing interval $[0, L]$ to determine the survival probability. For a particle initially at $x = x_0$, the concentration at time $t > 0$ is given by the Fourier series inversion

$$c(x, t=0) = \delta(x - x_0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right),$$

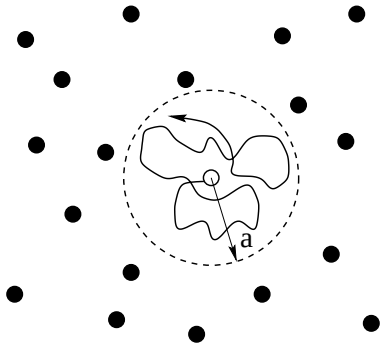


Figure 9.9: A configuration of traps (filled circles) and the trajectory of a diffusing particle. Also shown is the trap-free circle of radius a which is centered about the initial particle position. The probability that the particle remains in this circle is a lower bound for the exact particle survival probability in this configuration of traps.

which gives

$$A_n = \frac{2}{L}\sin\left(\frac{n\pi x_0}{L}\right).$$

Therefore the concentration within the interval is

$$c_L(x,t|x_0) = \frac{2}{L}\sum_{n=1}^{\infty}\sin\left(\frac{n\pi x}{L}\right)\,\sin\left(\frac{n\pi x_0}{L}\right)\;e^{-\left(\frac{n\pi}{L}\right)^2 Dt}. \tag{9.61}$$

For a fixed-length interval, we compute the survival probability by averaging over all initial particle positions and also integrating over all x. This gives

$$\begin{aligned}\overline{S_L(t)} &= \frac{1}{L}\int_0^L\int_0^L c_L(x,t|x_0)\,dx\,dx_0\\ &= \frac{8}{\pi^2}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^2}\;e^{-\frac{(2m+1)^2\pi^2}{L^2}Dt}.\end{aligned} \tag{9.62}$$

Next, we obtain the configuration-averaged survival probability by averaging this expression over the distribution of lengths of trap-free intervals. The simplest and most natural situation is a random distribution of traps at density ρ, for which the interval-length distribution is $P(L) = \rho e^{-\rho L}$. This gives the formal solution for the average survival probability

$$\begin{aligned}\langle S(t)\rangle &\equiv \langle\overline{S_L(t)}\rangle\\ &= \frac{8\rho}{\pi^2}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^2}\int_0^{\infty}\;e^{-\frac{(2m+1)^2\pi^2}{L^2}Dt}\;e^{-\rho L}\,dL.\end{aligned} \tag{9.63}$$

This integral has very different short- and long-time behaviors. In the former case, intervals of all lengths contribute to the survival probability, while at long times optimal-length intervals give the main contribution to the survival probability. This latter behavior is not visible until the survival probability has decayed to a vanishingly small and experimentally-unattainable value. In fact, the best strategy to observe the long-time behavior (by simulation) is to consider a system with a high concentration of traps.

Long-time behavior

In the long-time limit, clearly the first term in the series for $\langle S(t)\rangle$ in Eq. (9.63) eventually dominates. If we retain only this term, it is relatively easy to determine the asymptotic behavior of the integral in Eq. (9.63).

As a function of L, the first exponential factor in this equation rapidly increases to 1 as $L \to \infty$, while the second exponential factor decays rapidly with L. Thus the integrand has a peak as a function of L which becomes progressively sharper as $t \to \infty$. We may therefore determine the asymptotic behavior of $\langle S(t)\rangle$ by the Laplace method.

To apply this method, we first rewrite Eq. (9.63) as $\langle S(t)\rangle \sim \int_0^\infty e^{f(L)}\, dL$, and then we fix the location of the maximum by defining the dimensionless length $\ell \equiv L/L^*$ to transform the integral to

$$\begin{aligned}\langle S(t)\rangle &= \frac{8\rho L^*}{\pi^2}\int_0^\infty \exp\left[-(\rho^2 Dt)^{1/3}[(\pi^2/2)^{2/3}\frac{1}{\ell^2} + (2\pi^2)^{1/3}\ell]\right]\, d\ell,\\ &\equiv \frac{8\rho L^*}{\pi^2}\int_0^\infty \exp\left[-(\rho^2 Dt)^{1/3}\, g(\ell)\right]\, d\ell. \end{aligned} \tag{9.64}$$

The integrand now has an increasingly sharp maximum at a fixed location as $t \to \infty$. We therefore expand $g(\ell)$ to second order about its maximum and perform the resulting Gaussian integral to obtain the leading behavior of $\langle S(t)\rangle$. From the condition that $g'(\ell^*) = 0$, we find $\ell^* = 1$, $g(\ell^*) = 3(\pi/2)^{2/3}$ and $g''(\ell^*) = -3 \times (2\pi^2)^{1/3}$. Therefore

$$\begin{aligned}\langle S(t)\rangle &= \frac{8\rho L^*}{\pi^2}\int_0^\infty \exp\left[-(\rho^2 Dt)^{1/3}\, g(\ell)\right]\\ &\sim \frac{8\rho l^*}{\pi^2}\int_0^\infty \exp\left[-(\rho^2 Dt)^{1/3}\,[g(\ell^*) + \frac{1}{2}(\ell-\ell^*)^2 g''(\ell^*)]\right]\\ &\sim \frac{8\rho L^*}{\pi^2}\sqrt{\frac{2\pi}{(\rho^2 Dt)^{1/3}|g''(\ell^*)|}}\exp\left[-(\rho^2 Dt)^{1/3}\, g(\ell*)\right]\\ &= \frac{8\times 2^{2/3}}{3^{1/2}\pi^{7/6}}(\rho^2 Dt)^{-1/6}\,\exp(-3(\pi^2\rho^2 Dt/4)^{1/3}). \end{aligned} \tag{9.65}$$

The basic feature of this result is the relatively slow $e^{-t^{1/3}}$ asymptotic decay of $\langle S(t)\rangle$ compared to the exponential decay for the survival probability in a fixed-length interval. This slower decay stems from the contribution of optimal intervals whose length ℓ^* grows as $t^{1/3}$. Although such large intervals are rare, their contribution to the survival probability is asymptotically dominant. In Subsection 9.6.0.1, we shall see how these extreme intervals are the basis for the Lifshitz tail argument which provides the asymptotic decay of $\langle S(t)\rangle$ in arbitrary spatial dimension. Finally, if one is interested in only the correct controlling factor in the asymptotic survival probability, one can merely evaluate $f(L)$ at its maximum of $L^*(t) = (2\pi^2 Dt/\rho)^{1/3}$ and then estimate $\langle S(t)\rangle$ as $e^{f(L^*)} \sim e^{-\text{const.}\times(\rho^2 Dt)^{1/3}}$.

9.6.0.1 Short-Time Behavior

It is instructive to study the short-time behavior of $\langle S(t)\rangle$, both because the time dependence is interesting and because this limit indicates that the crossover to the asymptotic behavior for $\langle S(t)\rangle$ is very slow. In fact, the asymptotic decay does not arise until the density has decayed to an extremely small value. Thus although there is considerable theoretical appeal in understanding the long-time decay of the trapping reaction, its practical implications are limited.

There are many ways to estimate the short-time behavior. One crude approach is to notice that, at early times, the factor e^{-Dt/L^2} reaches 1 as a function of L (at $L \approx \sqrt{Dt}$) before there is an appreciable decay in the factor $e^{-\rho L}$ (at $L \approx 1/\rho$). Thus to estimate $\langle S(t)\rangle$, we may cut off the lower limit of the integral at $\sqrt{Dt}$ and replace the factor e^{-Dt/L^2} by 1. Using this approximation, the time dependence of the survival probability is

$$\begin{aligned}\langle S(t)\rangle &\approx \int_{\sqrt{Dt}}^\infty e^{-\rho L}\, dL\\ &\approx e^{-\text{const.}\times\rho\sqrt{Dt}}. \end{aligned} \tag{9.66}$$

This short-time behavior extends until $t \sim 1/(D\rho^2)$, which translates to the diffusion distance being of the order of the mean separation between traps.

A more rigorous approach is to use the fact we should keep all the series terms in Eq. (9.63). As shown in Weiss' book, this series can be evaluated easily by defining $\epsilon = \pi^2 Dt/L^2$ and noting that $dS/d\epsilon$ has the form

$$\left\langle \frac{\partial S(t)}{\partial \epsilon} \right\rangle = \frac{8\rho}{\pi^2} \int_0^\infty \left(\sum_{m=0}^{\infty} e^{-(2m+1)^2\epsilon} \right) e^{-\rho L}\, dL.$$

We can estimate the sum by replacing it with an integral, and then we can easily perform the average over L, with the result

$$\langle S(t) \rangle \sim e^{-\rho\sqrt{8Dt/\pi}}. \tag{9.67}$$

Now we may roughly estimate the crossover between the short- and the long-time limits by equating the exponents in Eqs. (9.65) and (9.67). This gives the numerical estimate $\rho^2\, Dt \approx 269$ for the crossover time. Substituting this into the above expression for the short-time survival probability shows that $\langle S(t) \rangle$ must decay to approximately 4×10^{-12} before the long-time behavior represents the main contribution to the survival probability. In fact, because of the similarity of the short- and the long-time functional forms, the crossover is very gradual, and one must wait much longer still before the asymptotic behavior is clearly visible. Although this discussion needs to be interpreted cautiously because of the neglect of the power-law factors in the full expressions for the survival probability, the basic result is that the asymptotic survival probability is of marginal experimental utility. In spite of this deficiency, the question about the asymptotic regime is of fundamental importance, and it helps clarify the role of exceptional configurations in determining the asymptotic survival probability.

Lifshitz Argument for General Spatial Dimension

In higher dimensions, it is not possible to perform this average directly. As a much simpler alternative, we will apply a Lifshitz argument to obtain the asymptotic behavior of the survival probability. Part of the reason for presenting this latter approach is its simplicity and wide range of applicability. One sobering aspect, however, is that the asymptotic survival probability does not emerge until the density has decayed to an astronomically small value. Such a pathology typically arises when a system is controlled by rare events. This serves as an important reality check for the practical relevance of the Lifshitz argument.

The Lifshitz approach has emerged as an extremely useful tool to determine asymptotic properties in many disordered and time-varying systems. If we are interested *only* in asymptotics, then it is often the case that a relatively small number of extreme configurations provide the main contribution to the asymptotics. The appeal of the Lifshitz approach is that these extreme configurations are often easy to identify and the problem is typically straightforward to solve on these configurations.

In the context of the trapping reaction, we first identify the large trap-free regions which give the asymptotically dominant contribution to the survival probability. Although such regions are rare, a particle in such a region has an anomalously long lifetime. By optimizing the survival probability with respect to these two competing attributes, we find that the linear dimension of these extreme regions grows as $(Dt)^{1/(d+2)}$ for isotropic diffusion, from which we can easily find the asymptotic survival probability.

9.6.0.2 Isotropic Diffusion

It is convenient to consider with a lattice system in which each site is occupied by a trap with probability p and in which a single particle performs a random walk on free sites. The average survival probability $\langle S(t) \rangle$ is obtained by determining the fraction of random-walk trajectories which do not hit any trap up to time t. This fraction must be averaged over all random-walk trajectories *and* over all trap configurations.

An important aspect of these averages is that they may be performed in either order, and it is more convenient to first perform the latter. For a given trajectory, each visited site must not be a trap for the particle to survive, while the state of the unvisited sites can be arbitrary. Consequently, a walk which has visited s *distinct* sites survives with probability q^s, with $q = (1 - p)$. Then the average survival probability is

$$\langle S(t) \rangle = z^{-N} \sum_s C(s,t) q^s \equiv \langle q^s \rangle, \tag{9.68}$$

where $C(s,t)$ is the number of random walks which visit s distinct sites at time t and z is the lattice coordination number. Notice that the survival probability is an exponential-order moment of the distribution of visited sites. It is this exponential character which leads to the anomalous time dependence of the survival probability.

Clearly the survival probability for each configuration of traps is bounded from below by the contribution which arises from the *largest* spherical trap-free region centered about the initial particle position (Fig. 9.9). This replacement of the configurational average by a simpler set of extremal configurations is the essence of the Lifshitz tail argument. The probability for such a region to occur is simply q^V, where $V = \Omega_d r^d$ is the number of sites in this d-dimensional sphere of radius r. We determine the probability for a particle to remain inside this sphere by solving the diffusion equation with an absorbing boundary at the sphere surface. This is a standard and readily-soluble problem, and the solution is merely outlined.

Since the system is spherically symmetric, we separate the variables as $c(r,t) = g(r)f(t)$ and then introduce $h(r) = r^\nu g(r)$, with $\nu = \frac{d}{2} - 1$ to transform the radial part of the diffusion equation into the Bessel differential equation

$$h''(x) + \frac{1}{x}h'(x) + \left(1 - \frac{1}{x^2}\left(\frac{d}{2} - 1\right)^2\right) h(x) = 0,$$

where $x = r\sqrt{k/D}$, the prime denotes differentiation with respect to x, and the boundary condition is $h(a\sqrt{k/D}) = 0$, where a is the radius of the trap-free sphere. Correspondingly $f(t)$ satisfies $\dot{f} = -kf$. In the long-time limit, the dominant contribution to the concentration arises from the slowest decaying mode in which the first zero of the Bessel function $J_{d/2}(r\sqrt{k/D})$ occurs at the boundary of the sphere. Thus the survival probability within a sphere of radius a asymptotically decays as

$$S(t) \propto \exp\left(-\frac{\mu_d^2 Dt}{a^2}\right),$$

where μ_d is the location of the first zero of the Bessel function in d dimensions.

To obtain the configuration-averaged survival probability, we average this survival probability for a fixed-size sphere over the radius distribution of trap-free spheres. This gives the lower bound for the average survival probability,

$$\langle S(t)\rangle_{\rm LB} \propto \int_0^\infty \exp\left[-\frac{\mu_d^2 Dt}{r^2} + \Omega_d r^d \ln q\right] r^{d-1}\, dr. \tag{9.69}$$

This integrand becomes sharply peaked as $t \to \infty$, and we can again estimate the integral by the Laplace method. As in one dimension, we rescale variables to fix the location of the maximum. Writing the integrand in Eq. (9.69) as $\exp[-F(r)]$ and differentiating with respect to r, we find that the maximum of F occurs at

$$r^* = \left(-\frac{2\mu_d^2 Dt}{\Omega_d d \ln q}\right)^{1/(d+2)}.$$

This defines the radius of the trap-free region which gives the dominant contribution to the survival probability at time t. We now rewrite F in terms of the scaled variable $u = r/r^*$ to give

$$F(u) = -(\mu_d^2 Dt)^{d/(d+2)}(-\Omega_d \ln q)^{2/(d+2)}\left[\left(\frac{d}{2u}\right)^{2/(d+2)} + \left(\frac{2u}{d}\right)^{d/(d+2)}\right].$$

We now evaluate the integral by expanding $F(u)$ to second order in u and performing the resulting Gaussian. This gives, for the controlling exponential factor in the average survival probability,

$$\begin{aligned}\langle S(t)\rangle_{\rm LB} &\propto \exp\left[-\text{const.} \times (Dt)^{d/(d+2)}(\ln w)^{2/(d+2)}\right] \\ &\equiv \exp[-(t/\tau)^{2/(d+2)}].\end{aligned} \tag{9.70}$$

There are two noteworthy points about this last result. First, this type of stretched exponential behavior is not derivable by a classical perturbative expansion, such as an expansion in the density of traps. Second,

as in the case of one dimension, the asymptotic decay in Eq. (9.70) again does not set in until the density has decayed to an astronomically small value. We can again obtain a rough estimate for this crossover time by comparing the asymptotic survival probability with the survival probability in the short-time limit. A cheap way to obtain the latter is to expand Eq. (9.68) as $\langle q^s\rangle = \langle 1 + s\ln q + (s\ln q)^2/2 + \ldots\rangle$, retain only the first two terms, and then re-exponentiate. This then gives

$$\langle S(t)\rangle_{\text{short time}} \approx q^{\langle s\rangle} \to e^{-\rho D t\, a^{d-2}}, \tag{9.71}$$

where a is the lattice spacing and we have assumed the limit of a small concentration of traps. By comparing the asymptotic form Eq. (9.70) with the short-time approximation of (9.71), we can infer the crossover time between short-time and asymptotic behavior and then the value of the survival probability at this crossover point. The detailed numerical evaluation of these numbers is tedious and unenlightening; however, the basic result is that the survival probability begins to show its asymptotic behavior only after it has decayed to a microscopically small value. In fact, the crossover to asymptotic behavior occurs earliest when the concentration of traps is large. This is counter to almost all simulation studies of the trapping reaction.

9.7 Spatially Dependent Aggregation

Aggregation often proceeds in environments that are spatially in-homogeneous. If the basic transport mechanism is diffusion, such a spatially dependent aggregation is governed by an infinite system of *partial* differential equations

$$\frac{\partial c_k}{\partial t} = D_k\Delta c_k + \frac{1}{2}\sum_{i+j=k} K_{ij}\, c_i\, c_j - c_k \sum_{j\geq 1} K_{kj}\, c_j \tag{9.72}$$

These **reaction-diffusion** equations are extremely complicated. Even for the simplest model with constant reaction rates and mass-independent diffusion coefficient,[1] is generally intractable.

Aggregation in a confined region with adsorption on the walls

For the model with constant reaction rates, $K_{ij} = 2K$, and mass-independent diffusion coefficients, $D_k = D$, the reaction-diffusion equations for the densities $c_k(t,\mathbf{r})$

$$\frac{\partial c_k}{\partial t} = D\Delta c_k + K\sum_{i+j=k} c_i\, c_j - 2K c_k N \tag{9.73}$$

should be solved inside the domain $\mathcal{D}$ subject to the initial condition

$$c_k(0,\mathbf{r}) = \delta_{k,1}\,, \qquad \mathbf{r}\in\mathcal{D} \tag{9.74}$$

and the adsorbing condition on the boundary $\partial\mathcal{D}$ of the domain

$$c_k(t,\mathbf{r}\in\partial\mathcal{D}) = 0\,. \tag{9.75}$$

The problem (9.73)–(9.75) is mathematically intractable even for simplest domains $\mathcal{D}$. Take for instance the cluster density. It satisfies a single reaction-diffusion equation

$$\frac{\partial N}{\partial t} = D\Delta N - KN^2, \quad N(t,\mathbf{r}\in\partial\mathcal{D}) = 0 \tag{9.76}$$

This nonlinear partial differential equation has not been solved. The behavior, however, is conceptually simple as both aggregation and adsorption are helping each other rather than competing, e.g. both processes reduce the number of clusters. To see which of the two effects dominates in the long time limit let us consider the behavior if one of the effect was absent. Disregarding adsorption (physically this would occur if boundaries

[1] Smoluchowsi's argument in the beginning of this chapter makes the model with constant reaction rates reasonable; a diffusion coefficient that does not decrease with mass is harder to justify.

do not absorb clusters) we recover the already known result $N = (1 + Kt)^{-1}$. Disregarding aggregation, we arrive at a linear, and therefore tractable problem for a diffusion equation. A (formal) solution reads

$$N(t, \mathbf{r}) = \sum_{n\geq 1} A_n \, e^{-\lambda_n Dt} \, \psi_n(\mathbf{r}) \tag{9.77}$$

where $0 < \lambda_1 < \lambda_2 \ldots$ are the eigenvalues of the Laplace operator with Dirichlet boundary conditions, $\psi_n(\mathbf{r})$ are the corresponding eigenfunctions

$$(\Delta + \lambda_n)\psi_n(\mathbf{r}) = 0, \qquad \psi_n(\mathbf{r} \in \partial\mathcal{D}) = 0$$

and the amplitudes A_n are fixed by the initial condition.

The exponential behavior is asymptotically much steeper than the power-law decay $N = (1+Kt)^{-1}$ characterizing the homogeneous situation. Hence adsorption is asymptotically more important than aggregation, and the cluster density eventually exhibits an exponential decay

$$N(t, \mathbf{r}) \sim e^{-\lambda_1 Dt} \, \psi_1(\mathbf{r}) \tag{9.78}$$

This behavior is valid when time exceeds the characteristic time scale of adsorption (which is essentially a time to diffuse across the domain $\mathcal{D}$)

$$t \gg t_{\rm ads} = \frac{1}{\lambda_1 D} \sim \frac{L^2}{D} \tag{9.79}$$

Here we used the estimate $\lambda_1 \sim L^{-2}$ expressing the smallest eigenvalue via the characteristic length scale L of the reaction domain $\mathcal{D}$. The characteristic time scale of aggregation is $t_{\rm agg} = K^{-1}$. The behavior of the cluster density depends on the relative magnitude of the characteristic time scales for adsorption and aggregation:

1. $t_{\rm ads} \ll t_{\rm agg}$. In this situation, aggregation is irrelevant; the series solution (9.77) is valid throughout the evolution.

2. $t_{\rm ads} \gg t_{\rm agg}$. In this situation, aggregation dominates in the intermediate time range $t \leq t_{\rm ads}$. More precisely, this is correct in the bulk of the reaction domain $\mathcal{D}$; in the boundary layer of width $\sqrt{Dt}$ near the domain boundary $\partial\mathcal{D}$, adsorption is important. Adsorption eventually wins when $t \geq t_{\rm ads}$, and the long-time behavior is given[2]

$$N(t, \mathbf{r}) \approx \frac{D}{KL^2} \, e^{-\lambda_1 Dt} \, \psi_1(\mathbf{r}) \tag{9.80}$$

Example 4. *Aggregation between two parallel absorbing plates.* Consider a domain $\mathcal{D}$ confined by two flat plates at $x = 0$ and $x = L$. Even in this situation the governing equations are unsolvable. The cluster density $N(t, x)$ obeys

$$\frac{\partial N}{\partial t} = D \, \frac{\partial^2 N}{\partial x^2} - KN^2 \tag{9.81}$$

with boundary conditions $N(t, 0) = N(t, L) = 0$. The eigenvalues are $\lambda_n = (\pi n/L)^2$, the corresponding eigenfunctions are $\psi_n = \sin(\pi nx/L)$, and the long time asymptotic is

$$N(t, x) \sim \exp\left(-\frac{\pi^2 Dt}{L^2}\right) \, \sin\left(\frac{\pi x}{L}\right)$$

Example 5. *Aggregation near an absorbing plate.* Let $x = 0$ be the absorbing flat plate. Clusters occupy the half-space $x > 0$ where they diffuse and aggregate. In this semi-infinite system, the characteristic time $t_{\rm ads}$ is infinite. Hence aggregation dominates far away from the plane, $x \gg \sqrt{Dt}$, and both aggregation and adsorption are relevant

[2] We estimated a prefactor in equation (9.78) by matching to the behavior at the end of the intermediate time range; this gives $(Kt_{\rm ads})^{-1} = t_{\rm agg}/t_{\rm ads} = D/KL^2$.

in the (growing) boundary layer with width of the order of $\sqrt{Dt}$. Thus asymptotically (when $t \gg t_{\rm agg} = K^{-1}$) the density $N(t,x)$ is expected to approach a scaling form

$$N(t,x) = (Kt)^{-1} f(\xi), \qquad \xi = \frac{x}{\sqrt{Dt}} \tag{9.82}$$

Using this scaling ansatz we reduce (9.81) to

$$f'' + \frac{1}{2}\,\xi f' + f(1-f) = 0 \tag{9.83}$$

The adsorbing boundary condition $N(t,0) = 0$ and the requirement that the scaling form (9.82) matches the bulk behavior $N(t, x\to\infty) = (Kt)^{-1}$ give

$$f(0) = 0, \qquad f(\infty) = 1 \tag{9.84}$$

While the ordinary differential equation (9.83) is much simpler than the original partial differential (9.81), it still does not admit a closed-form solution. The problem (9.83)–(9.84) must be solved numerically.

The cluster mass distribution $c_k(t,x)$ is the function of three variables — the discrete variable k and two continuous variables t, x. We anticipate that the mass distribution attains a scaling form

$$c_k(t,x) = (Kt)^{-2}\Phi(\xi,\eta) \tag{9.85}$$

in the scaling limit $k, t, x \to \infty$ with

$$\xi = \frac{x}{\sqrt{Dt}} = \text{finite}, \qquad \eta = \frac{k}{Kt} = \text{finite}$$

Inserting (9.85) into (9.73) we see that the scaling function $\Phi(\xi,\eta)$ obeys

$$\frac{\partial^2\Phi}{\partial\xi^2} + \frac{1}{2}\,\xi\,\frac{\partial\Phi}{\partial\xi} + \eta\,\frac{\partial\Phi}{\partial\eta} + \int_0^\eta d\eta'\,\Phi(\xi,\eta')\Phi(\xi,\eta-\eta') = 0$$

The boundary conditions for the scaling function are

$$\Phi(0,\eta) = 0, \qquad \Phi(\infty,\eta) = e^{-\eta}$$

Aggregation with a localized source

A spatially *localized* input leads to in-homogeneous aggregation. For the model with constant reaction and diffusion rates, the governing reaction-diffusion equations are

$$\frac{\partial c_k}{\partial t} = D\Delta c_k + K\sum_{i+j=k} c_i\,c_j - 2Kc_kN + J\delta_{k,1}\delta(\mathbf{r}) \tag{9.86}$$

where we considered a monomer input with strength J.

In the following we describe the behavior in three particular cases:

1. The monomers are injected uniformly on the plane $x = 0$. This is effectively a one-dimensional situation, so that the delta function in (9.86) becomes $\delta(x)$ and the Laplace operator is $\Delta = \partial^2/\partial x^2$.

2. The monomers are injected uniformly along the line $x = y = 0$. This is effectively a two-dimensional situation. The Laplace operator is

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\,\frac{\partial}{\partial r}$$

   due to cylindrical symmetry (here $r = \sqrt{x^2+y^2}$).

3. The monomers are injected at the origin $x = y = z = 0$. In this case the Laplace operator is

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\,\frac{\partial}{\partial r}$$

   where $r = \sqrt{x^2+y^2+z^2}$ and the second formula utilizes spherical symmetry.

Whenever possible, we consider all three cases together and distinguish them by the dimensionality $d = 1, 2, 3$. The actual system is *three-dimensional*, the dimensionality d just counts the number of directions along which the densities change.

The interplay between input and aggregation results in *stationary* limits for the cluster densities. Not all quantities, however, become stationary in the $t \to \infty$ limit.

Moments

The mass density $M(t, \mathbf{r})$ satisfies the diffusion equation with a source

$$\frac{\partial M}{\partial t} = D\Delta M + J\delta(\mathbf{r}) \tag{9.87}$$

and therefore (for initially empty system)

$$M(t, \mathbf{r}) = J \int_0^t \frac{d\tau}{(4\pi D\tau)^{d/2}}\, e^{-r^2/4D\tau} \tag{9.88}$$

Surprisingly, $M(t, \mathbf{r})$ reaches a stationary limit only when $d = 3$. In this case, the stationary mass density is a fundamental solution of the Laplace equation, $\Delta M = -(J/D)\delta(\mathbf{r})$, and hence $M = (J/D)(4\pi r)^{-1}$.

Overall, equation (9.88) leads to the following asymptotic behaviors (valid when $t \to \infty$ and $r \ll \sqrt{Dt}$):

$$M(t, \mathbf{r}) = \frac{J}{4\pi D} \times \begin{cases} \sqrt{\pi Dt} & d = 1 \\ \ln(4Dt/r^2) - \gamma + O(r^2/4Dt) & d = 2 \\ 1/r & d = 3 \end{cases} \tag{9.89}$$

where $\gamma = 0.577215\ldots$ is Euler's constant.

The cluster density $N(t, \mathbf{r})$ does reach a stationary limit. In this limit, the reaction-diffusion equation becomes

$$D\Delta N - KN^2 = 0 \tag{9.90}$$

An algebraic ansatz $N = Ar^{-n}$ solves (9.90) when the exponent $n = 2$ and $A = (8-2d)D/K$. This algebraic solution has an incorrect behavior near the origin, and the full solution is known (see below) only for $d = 1$ when the non-linear differential equation (9.90) has constant coefficients. The algebraic solution, however, provides the correct large distance asymptotic:

$$N \simeq \frac{(8-2d)D}{K}\, r^{-2} \tag{9.91}$$

Stationary mass distribution when $d = 1$

The problem simplifies when the monomers are injected on the plate $x = 0$ and the densities depend on one spatial variable x. Equation (9.90) becomes

$$DN'' - KN^2 + J\delta(x) = 0 \tag{9.92}$$

where $N'' = d^2N/dx^2$. The source vanishes when $x \neq 0$, and there equation (9.92) admits the integral of motion

$$D(N')^2 - \frac{2}{3}\, KN^3 = 0 \tag{9.93}$$

where the constant on the right-hand side is zero since $N \to 0$ as $|x| \to \infty$. Integrating (9.93) we find

$$N = \frac{D}{K}\, \frac{6}{(|x| + x_0)^2} \tag{9.94}$$

To determine the integration constant x_0 we integrate (9.92) over the tiny region $(-\epsilon, \epsilon)$ around the origin, and send $\epsilon \to 0$. We get

$$D[N'(+0) - N'(-0)] + J = 0$$

and using (9.94) we obtain

$$x_0 = \left(24\,\frac{D^2}{JK}\right)^{1/3}$$

To find the stationary densities we must solve

$$Dc_k'' + K\sum_{i+j=k} c_i\,c_j - 2Kc_kN + J\delta_{k,1}\delta(x) = 0 \tag{9.95}$$

The generating function $\mathcal{C}(x,z) = \sum_{k\geq 1} c_k(x)z^k$ satisfies

$$D\mathcal{C}'' + K(\mathcal{C}^2 - 2\mathcal{C}N) + Jz\delta(x) = 0$$

and therefore

$$D(\mathcal{C}-N)'' + K(\mathcal{C}-N)^2 - J(1-z)\delta(x) = 0$$

This equation is solved repeating the steps used in solving equation (9.92). Thus we arrive at

$$\mathcal{C}(x,z) = \frac{6D}{Kx_0^2}\left\{\frac{1}{[|\rho|+1]^2} - \frac{1}{[|\rho|+(1-z)^{-1/3}]^2}\right\}, \qquad \rho = \frac{x}{x_0} \tag{9.96}$$

The mass distribution has a simple form at the origin. We have

$$\mathcal{C}(0,z) = \frac{6D}{Kx_0^2}\left\{1-(1-z)^{2/3}\right\}$$

and expanding the generating function in a power series in z we obtain

$$c_k(0) = \frac{4D}{Kx_0^2}\,\frac{\Gamma\left(k-\frac{2}{3}\right)}{\Gamma\left(\frac{1}{3}\right)\,\Gamma(k+1)} \tag{9.97}$$

Far away from the source ($\rho \gg 1$) we have

$$\mathcal{C}(x,z) = \frac{12D}{Kx_0^2}\,\rho^{-3}\left\{(1-z)^{-1/3}-1\right\}$$

and therefore

$$c_k(x) = \frac{12D}{Kx_0^2}\,\rho^{-3}\,\frac{\Gamma\left(k+\frac{1}{3}\right)}{\Gamma\left(\frac{1}{3}\right)\,\Gamma(k+1)} \tag{9.98}$$

when $\rho \gg 1$.

Let us take the limit $\rho \to \infty$ and $k \to \infty$ in such a way that the scaling variable $\xi = k/\rho^3$ remains finite. Then the mass distribution attains the scaling form

$$c_k(x) = \frac{D}{Kx_0^2}\,\rho^{-5}\,\Phi(\xi)\,, \qquad \xi = \frac{k}{\rho^3} \tag{9.99}$$

Writing $1-z = s/\rho^3$ we recast (9.96) to

$$\mathcal{C}(x,z) = \frac{6D}{Kx_0^2}\,\rho^{-2}\left\{1-\left(1+s^{-1/3}\right)^{-2}\right\}$$

while using (9.99) and $z^k \simeq e^{-\xi s}$ and replacing summation by integration we get

$$\mathcal{C}(x,z) = \sum_{k\geq 1} c_k(x)z^k \simeq \frac{D}{Kx_0^2}\,\rho^{-2}\int_0^\infty d\xi\,\Phi(\xi)\,e^{-\xi s}$$

Therefore we found the Laplace transform of the scaling function

$$\int_0^\infty d\xi\,\Phi(\xi)\,e^{-\xi s} = 6\left\{1-\left(1+s^{-1/3}\right)^{-2}\right\} \tag{9.100}$$

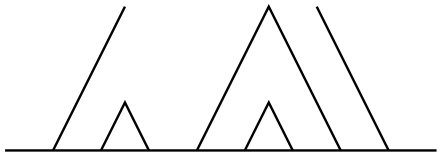


Figure 9.10: Ballistic annihilation with two-velocities.

9.8 Ballistic Annihilation

In ballistic annihilation, particles move at constant velocity and annihilation occurs whenever two particles meet.

We start with the case of bimodal velocity distributions. The two velocities can be taken to be equal in magnitude and opposite in sign: v_0 and $-v_0$. The concentration of the two particles are equal c_0.

The two-velocity problem is analytically tractable because it maps directly to the survival probability of a random walk in the presence of a trap. As in the traffic problem, a positive velocity particle is of course affected only by particles ahead of it. Its collision partner depends only on the velocities of the particles ahead but not on their actual positions. For the velocity configuration be $++-++---\cdots$, the 0th particle is bound to collide with the 7th particle (Fig.9.10). In general, it collides with the kth particle when the velocity sum $\sum_{i=0}^{m} v_i \geq 0$ for all $m < 2k$ but $\sum_{i=0}^{2k+1} v_i < 0$.

Let p_k the probability that $2k$ consecutive particles all annihilate among themselves. Manually, we find $p_0 = 1$, $p_1 = 1/4$ and $p_2 = 1/8$. This probability satisfies the recursion relation

$$p_k = \frac{1}{4}\sum_{j=1}^{k-1} p_j p_{k-1-j} \tag{9.101}$$

for $k > 0$ with $p_0 = 1$. The generating function $p(z) = \sum_{k=0}^{\infty} p_k z^k$ satisfies $\frac{z}{4}P^2(z) - P(z) + 1 = 0$. Its solution, $p(z) = 1 - \sqrt{1-z}$ yields the probabilities

$$p_k = 4^{-k}\frac{(2k)!}{k!(k+1)!}. \tag{9.102}$$

These probabilities allow calculation of the density of remaining particles for arbitrary spatial distributions. For simplicity, we consider a regular array of particles, with spacing all equal to $1/c_0$. For a positive particle to survive to time t it must be destined to collide with a particle of index $k > c_0 v_0 t$. Therefore, the particle concentration equals $c(t) = c_0 \sum_{k>c_0 v_0 t} p_k$. Using $p_k \sim k^{-3/2}$ leads to the concentration decay

$$c(t) \sim \left(\frac{c_0}{v_0 t}\right)^{1/2}. \tag{9.103}$$

Thus, the concentration decays algebraically with time, much slower compared with the exponential decay for traffic with bimodal velocity distributions. Moreover, the concentration depends on the initial condition (there is an explicit dependence on the initial concentration) in contrast with the single-species annihilation decay (9.14).

The long time behavior is dominated by fluctuations in the initial conditions, a behavior that is in some sense similar to both the traffic problem and the two-species annihilation reaction in low spatial dimensions. We employ the finite-size scaling argument (see box in chapter 7). In a finite system, there are initially N_+ and N_- particles with $N_+ + N_- = c_0 L$. The fluctuations in the particle numbers are characterized by the number difference $\Delta N = |N_+ - N_-|$ and since the initial concentrations are the same, this fluctuation grows diffusively with the total particle number $\Delta N \sim N^{1/2}$. The final state consists of all the excess majority particles, so $c(L) \sim \Delta/L \sim (c_0/L)^{-1/2}$. Since the only dynamical scale in the problem is the ballistic scale $v_0 t$ we anticipate that the time dependent concentration obeys the scaling relation $c(L,t) \sim (c_0 L)^{-1/2}\Phi(tv_0/L)$. In the infinite system size limit, the concentration should depend on time alone so $\Phi(z) \sim z^{-1/2}$ as $z \to 0$ therefore reproducing (9.103).

A rich behavior occurs when there are three types of velocities. In the symmetric case of three velocities $-v_0$, 0, and v_0 with equal concentrations of mobile particles, $c_+(0) = c_-(0)$ and $c_0(0) = 1 - c_-(0) - c_+(0)$, there is a phase transition at $c_0(0) = 1/4$. The problem reduces to the two-velocity case when $c_0(0) < 1/4$ with the mobile concentration decaying as $c_\pm(t) \sim t^{-1/2}$ and the immobile concentration decaying as $c_- \sim t^{-1}$. At the critical point, all the concentrations decay asymptotically as $t^{-2/3}$. Above the critical concentration, a finite fraction of the immobile particles survive, and the mobile particle concentration decays exponentially with time.

For continuous velocity distributions, the behavior is qualitatively similar to traffic flows. The velocity decays as $v \sim t^{-\beta}$ and the concentration as $c \sim t^{-\alpha}$. From dimensional analysis the exponent relation $\alpha + \beta = 1$ holds. The exponents vary continuously with the parameter μ in **an undefined equation that was labeled tf-piv**. The exponents values $\beta(\mu)$ differ from the traffic case, but qualitatively, they do exhibit a similar dependence on μ. Overall, the Boltzmann equation

$$\frac{\partial P(v,t)}{\partial t} = -P(v,t)\int_{-\infty}^{\infty} dv' |v - v'| P(v',t) \tag{9.104}$$

provides a decent approximation. For example, it predicts $\beta(0) = 0.230472$ compared with Monte Carlo simulation results $\beta = 0.196$.

Combining the results for traffic flows, ballistic agglomeration, and ballistic annihilation, we conclude that reaction processes with a ballistic transport are much less robust compared with their diffusive counterparts. Dimensional analysis is generally inappropriate for describing the behavior as exponents can be transcendental. Conservations laws play an important role. The notion of universality classes is also not too useful. While the most complete knowledge was obtained for the exactly solvable traffic model, exact solutions are very difficult and generally require different techniques for different problems.

Problems

Section 9.1

1. Following the approach of this section, determine the probability distribution for the number of particles when there are initially $n_0 > 1$ particles in the system.

Chapter 10

COMPLEX NETWORKS

The study of complex networks represents a non-traditional application of non-equilibrium statistical physics. As we shall discuss, the tools of the field seem particularly appropriate to quantify basic properties of complex networks, such as percolation transitions and many geometrical properties. In this chapter, we will present some of the simplest complex network models and apply the master equation to quantify many of their basic features through a dynamical approach.

10.1 Erdös-Rényi Random Graph

A simple and classic starting example is the *Erdös-Rényi (ER) random graph*. This graph consists of N nodes in which each node pair may be joined by a link according to a connection probability that we define as p/N, with $0 \leq p \leq N$. Since any pair of nodes may be connected equiprobably, the ER graph has no spatial structure. A striking feature of the ER graph is the existence of a phase transition, in the limit $N \to \infty$, at $p = 1$. For $p < 1$, a finite network consists of disconnected clusters whose a maximum size is of the order of $\ln N$. At $p = 1$ the size of the largest cluster beccomes of the order of $N^{2/3}$. In the $N \to \infty$ limit, this largest cluster becomes the incipient infinite cluster; the term incipient refers to the fact that this cluster comprises a vanishing fraction of all nodes as $N \to \infty$. Finally, for $p > 1$ the largest cluster consists of a finite fraction of all nodes. For $p = N$, all $N(N-1)/2$ pairs of nodes are connected to give the *complete graph*.

While the ER graph is usually defined as a static problem — each link is independently present with probability p/N — we recast the ER graph *dynamically* by creating links between nodes at a constant rate. Within this formulation, the master equation provides in a simple way to determine the structure of the ER graph. A similar dynamical perspective was used to determine the coverage evolution in irreversible adsorption (Chapter 6), from which the final coverage — an ostensibly static quantity — emerged as a simple byproduct.

mention ER graph is like a tree with $z^g \approx N$ generations.

Degree distribution

We build the ER graph by starting with N disconnected nodes and then introducing links one by one between randomly-selected node pairs. The two nodes may be the same, and also, more than one link may be created between a pair of nodes. However, these two processes occur with a vanishingly small probability when $N \to \infty$ and may be ignored. For convenience, we define the rate at which each link is introduced as $(2N)^{-1}$. The total number of links at time t is then $Nt/2$ and the average degree equals t. Here degree is the number of links that are attached to a node. Thus the average degree evolves by a stochastic process in which $k \to k+1$ at rate 1.

As we shall see, the distribution of degrees is an important characteristic of complex networks. We define the degree distribution as n_k, the fraction of nodes of degree k. Nodes of degree k are created from nodes of degree $k-1$ at rate 1, and nodes of degree k are also lost at rate 1 due to the creation of nodes of degree

$k+1$. The degree distribution therefore satisfies the master equation of the Poisson process

$$\frac{dn_k}{dt} = n_{k-1} - n_k \,, \tag{10.1}$$

which applies for all $k \geq 0$ if we impose the additional condition $n_{-1} \equiv 0$. For a network with no links at $t = 0$, the initial condition is $n_k(0) = \delta_{k,0}$. Eqs. (10.1) may then be solved one by one starting with n_0 (see problem 10.1), and the degree distribution is

$$n_k = \frac{t^k}{k!}\, e^{-kt} \ . \tag{10.2}$$

From this expression, the mean degree equals the time, $\langle k \rangle = t$, while the standard deviation $\sqrt{\langle k^2 \rangle - \langle k \rangle^2} = \sqrt{t}$. Thus the degree distribution becomes sharp in the thermodynamic limit.

The percolation transition

We probe the percolation transition in the ER graph by studying the cluster size distribution. Here a cluster is defined as the set of all nodes that are connected by links into a single connected component. Initially the network consists of N clusters of size 1. As links are added, clusters merge and their number systematically decreases while their mean size grows. Since a link occurs equiprobably between any pair of nodes, there are $i \times j$ ways to join disconnected clusters of sizes i and j to create a cluster of size $i + j$; consequently, the overall rate for this event is $ij/(2N)$. This process is precisely product kernel aggregation discussed in chapter 4, and we can make use of the results derived therein to determine the cluster size distribution.

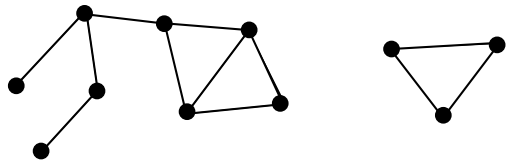


Figure 10.1: A realization of the ER graph that consists of two clusters: one of size 8 and one of size 3.

Let $c_k(t)$ be the density of clusters that contain k nodes at time t. The cluster size distribution obeys the master equation

$$\frac{dc_k}{dt} = \frac{1}{2} \sum_{i+j=k} (ic_i)(jc_j) - k\, c_k, \tag{10.3}$$

with $c_k(0) = \delta_{k,1}$ for the disconnected initial condition. The gain term accounts for the merger between two clusters whose sizes sum to k, and the loss term accounts for the loss of clusters of size k due to their linking with other clusters. As a preliminary, it is useful to study moments of the size distribution, $M_n = \sum_k k^n c_k$. The first moment, M_1, is just the fraction of nodes that belong to finite clusters; this quantity therefore equals 1 as long as there is no gelation, a condition that holds for $t < 1$ (see Sec. 4.1 for a detailed discussion of this point). As derived in Eq. (4.30), the second moment, which gives the mean cluster size, obeys the rate equation $\dot{M}_2 = M_2^2$ for $t < 1$. With the initial condition $M_2(0) = 1$, the solution is simply

$$M_2 = (1 - t)^{-1}, \tag{10.4}$$

which shows that an infinite cluster forms at $t = 1$ when the number of nodes of the ER graph is infinite. As t increases beyond the percolation point, this infinite cluster contains a finite fraction of all nodes and eventually engulfs the entire system.

From our earlier discussion of product-kernel aggregation, the cluster size distribution is [see Eq. (4.38)]

$$c_k(t) = \frac{k^{k-2}}{k!}\, t^{k-1}\, e^{-kt} \ . \tag{10.5}$$

While this exact distribution seems to have a smooth time dependence, there is a dramatic change in behavior as t passes through 1 that can be seen by using Stirling's approximation to give the asymptotic behaviors:

$$c_k(t) \sim \begin{cases} e^{-k(t-\ln t-1)} & t<1; \\ (2\pi)^{-1/2}k^{-5/2} & t=1. \end{cases} \tag{10.6}$$

The existence of a power-law distribution at $t = 1$ signals the percolation transition where the mean cluster size diverges.

What happens when the ER graph is finite? The transition is no longer sharp, and singular behavior is replaced by finite-size scaling laws. For example, for a finite network, the incipient infinite cluster now becomes just the largest cluster. Its size, M, can be estimated by the extremal criterion,

$$N\sum_{k\geq M}^{\infty} c_k = 1, \tag{10.7}$$

which states that there should be a single cluster whose size is in the range $[M, \infty]$. Using the asymptotic forms in Eqs. (10.6) for $c_k(t)$ in the extremal criterion, and approximating the sum by an integral we obtain

$$M \sim \begin{cases} \ln N & t<1; \\ N^{2/3} & t=1. \end{cases} \tag{10.8}$$

Thus clusters are at most of size $\ln N$ below percolation, while a "giant" cluster of size $N^{2/3}$ emerges as percolation is approached. When does this cluster appear? Close to the percolation time, we may estimate the typical cluster mass as

$$M_2 \approx \sum^{M} k^2 c_k \sim \int^{M} k^2\ k^{-5/2}\, dk \sim M^{1/2},$$

and equating this result to $M_2 = (1-t)^{-1}$ in Eq. (10.4), we obtain $M \sim (1-t)^{-2}$. Since M also scales as $N^{2/3}$, we obtain the percolation time in a finite network:

$$t \sim 1 - N^{-1/3}. \tag{10.9}$$

Paths and cycles

A deeper characterization of the ER graph may be obtained by studying paths and cycles in the network. A pair of nodes that are connected by a sequence of links forms a *path*. How do paths evolve with time? When a newly-added link connects the ends of two paths of lengths n and m, the result is a path of length $n + m + 1$. For $\ell > 0$, the density of *distinct* paths that contain l links at time t, $P_\ell(t)$, evolves as

$$\frac{dP_\ell}{dt} = \sum_{n+m=\ell-1} P_n P_m. \tag{10.10}$$

The initial condition is $P_\ell(0) = \delta_{l,0}$ and we define $P_0(t) = 1$. The solution of (10.10) is simply

$$P_\ell = t^\ell. \tag{10.11}$$

For example, $P_1 = t$ corresponds to the link density being equal to $t/2$ and that every link corresponds to two distinct paths of length 1 **ambiguous**. From this path length distribution, the total density of paths and the typical path length,

$$\begin{aligned} P_{\rm tot} &\equiv \sum_\ell P_\ell = \frac{1}{1-t}\ , \\ \langle \ell \rangle &= \frac{\sum_\ell \ell P_\ell}{\sum_\ell P_\ell} = \frac{t}{1-t}\ , \end{aligned} \tag{10.12}$$

respectively, both diverge at $t = 1$.

When a link directly joins two nodes that are already on the same path, a cycle forms. Let the average number of cycles of length ℓ at time t be $Q_\ell(t)$. This quantity is coupled to the path length density through the rate equation

$$\frac{dQ_\ell}{dt} = \frac{1}{2}\, P_{\ell-1}. \tag{10.13}$$

The right-hand side equals the link creation rate $1/(2N)$ times the total number of paths $NP_{\ell-1}$. Solving this equation, the cycle length distribution is

$$Q_\ell = \frac{t^\ell}{2\ell}. \tag{10.14}$$

Consequently, the total number of cycles in the system is $Q_{\rm tot} \equiv \sum_\ell Q_\ell = \frac{1}{2}\ln\frac{1}{1-t}$, which diverges weakly as $t \to 1$.

10.2 Sequentially Growing Networks

Sequential growth describes the evolution of many networked systems, such as the Internet and the world-wide web, where new routers or websites are added incrementally. It is natural to model such growing networks, by adding nodes one by one with each new node attaching to a "target" node, or a set of target nodes, with attachment rate A_k that depends only on the degree of the target (Fig. 10.2). The number of nodes N therefore plays the role of a time-like variable and we sometimes refer to N as the "time". The case where the attachment rate A_k increases with k defines *preferential attachment*, which encapsulates the notion of the "rich get richer". For example, in the context of scientific citations, preferential attachment means that a currently well-cited paper is more likely to be well cited in the future simply by virtue of being well cited now. We now apply the master equation to elucidate the structure of such growing networks.

Uniformly Growing Tree

As a starting example, we study the simpler case of the *uniformly growing tree* (UGT), also known as the random recursive tree. The growth rules of a UGT at any stage of its evolution are:

1. Pick one of the nodes of the UGT with uniform probability.
2. Introduce a new node that links to the target node.

Starting with the initial state of a single node, these steps repeated until the tree reaches a desired number of nodes N. Each node is distinguished by the order in which it is introduced so that there are $N!$ distinct trees of N nodes. Since each newly-introduced node has a single link, no closed loops can be generated. Thus if the graph initially is a tree, it remains a tree.

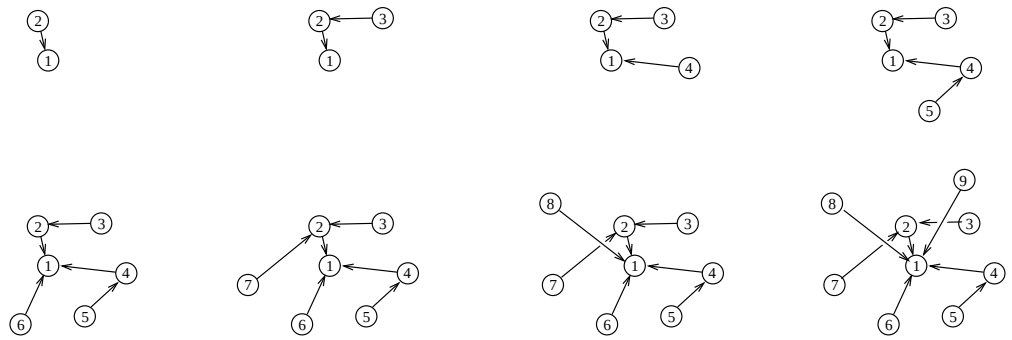


Figure 10.2: Evolution of one realization of a uniformly growing tree (upper left to lower right).

The degree distribution

Let's determine the degree distribution of the UGT, namely, the number of nodes of degree k when the network contains N nodes, $N_k(N)$. This distribution is distinct for each realization of the UGT, and, from the statistical physics perspective, the interesting quantity is degree distribution *averaged* over all realizations of trees with fixed N. Although the average distribution does not have a simple form for small N, it simplifies considerably when N is large. In this limit, the average number of nodes of degree k evolves according to

$$\frac{dN_k}{dN} = \frac{N_{k-1} - N_k}{N} + \delta_{k1}, \tag{10.15}$$

where we denote the average degree distribution by N_k **average notation**. This master equation is essentially the same as that for the ER graph, Eq. (10.1), except for the additional delta-function term that accounts for the single outgoing link of the new node.

To get a feeling for the solution, let's start by solving the master equations (10.15) one by one. With the understanding that $N_{-1} = 0$, the master equations are, explicitly:

$$\begin{aligned}
\dot{N}_0 &= -\frac{N_0}{N} \\
\dot{N}_1 &= \frac{N_0 - N_1}{N} + 1 \\
\dot{N}_2 &= \frac{N_1 - N_2}{N} \\
\dot{N}_3 &= \frac{N_2 - N_3}{N},
\end{aligned}$$

etc., where the overdot denotes differentiation with respect to N. The solution to the first equation is simply $N_0 = 1/N$. We now rewrite the equation for N_1 as $(\dot{N_1 N}) = N + N_0$, with asymptotic solution $N_1 \sim N/2$. By the same method, the equation for N_2 becomes $(\dot{N_2 N}) = N_1$, from which $N_2 \sim N/4$. This pattern of behavior continues for all k so that we conclude that all the N_k are proportional to N.

It therefore is convenient to work with the density of nodes of degree k, $n_k \equiv N_k/N$. In terms of this quantity, Eq. (10.15) reduces to

$$n_k = n_{k-1} - n_k + \delta_{k1}, \tag{10.16}$$

which are trivially soluble. Starting with $n_0 = 0$, we obtain $n_1 = \frac{1}{2}$, $n_2 = \frac{1}{4}$, *etc.*, and the general solution is simply $n_k = 2^{-k}$. Thus the UGT has a rapidly decaying degree distribution in which the average degree equals 2 and the largest degree is of order $\ln N$ for a network of N nodes.

Genealogical tree

It is revealing to study the genealogy underlying a UGT. We build this genealogy by taking generation $g = 0$ to be the initial node. Nodes that attach to those in generation g form generation $g + 1$. For example, in the final network of Fig. 10.2, node 1 is the "ancestor" of 2, while nodes 3 and 7 are the "descendants" of 2. There are 5 nodes in generation $g = 1$ and 3 in $g = 2$, leading to the genealogy of Fig. 10.3.

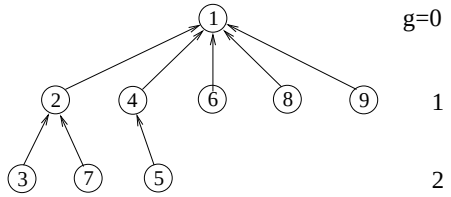


Figure 10.3: Genealogy of the network in Fig. 10.2 with nodes arranged according to generation number. The node indices indicate when each is introduced.

How many generations are there in a tree of N nodes? What is the size of the $g^{\rm th}$ generation, $L_g(N)$? To determine $L_g(N)$, note that $L_g(N)$ increases when a new node attaches to a node in generation $g-1$, an event that occurs with probability L_{g-1}/N. This gives the evolution equation $\dot{L}_g(N) = L_{g-1}/N$, with solution $L_g(\tau) = \tau^g/g!$, where $\tau = \ln N$. Using Stirling's approximation, we see that the generation size $L_g(N)$ therefore grows with g, when $g < \tau$, and then decreases and becomes of order 1 when $g = e\tau$. The genealogical tree therefore contains approximately $e\tau$ generations for a tree of N nodes. This latter result also determines the diameter of the tree, since the diameter (also the maximum distance between any pair of nodes) is twice the distance from the root to the last generation. Therefore the diameter of the tree scales as $2e\tau \approx 2e \ln N$; this is the same dependence on N as in the Erdös-Rényi random graph.

Redirection

We now generalize the UGT to incorporate the mechanism of *redirection*. In redirection a new node **n** is introduced and an earlier node **x** is uniformly selected as a target. With probability $1-r$, the link from **n** to **x** is created. However, with probability r, the link is redirected to the ancestor **y** of node **x** (Fig. 10.4).

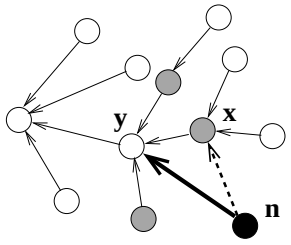


Figure 10.4: Illustration of redirection. The new node (solid) selects a target node **x** uniformly at random. With probability $1-r$ a link is established to this target (dashed arrow), while with probability r the link is established to **y**, the ancestor of **x** (thick solid arrow). The rate at which attachment to **y** occurs by redirection is proportional to the number of its upstream neighbors (shaded).

According to the defining processes of redirection shown in Fig. 10.4, the degree distribution $N_k(N)$ evolves according to

$$\frac{dN_k}{dN} = \frac{1-r}{N}\left[N_{k-1} - N_k\right] + \delta_{k1} + \frac{r}{N}\left[(k-2)N_{k-1} - (k-1)N_k\right]. \tag{10.17}$$

The first three terms correspond to the growth processes of the UGT, whose master equation (10.15) is recovered for redirection probability $r = 0$. The last two terms account for the change in N_k due to redirection. To understand their origin, consider the gain term. Since the initial node is chosen uniformly, if redirection does occur, the probability that a node of degree $k-1$ receives the newly-redirected link is proportional to the number of its incoming links, which equals $k-2$ (shaded nodes in Fig. 10.4). A similar argument applies for the redirection-driven loss term.

Combining the terms in Eq. (10.17), the master equation becomes

$$\frac{dN_k}{dN} = \frac{r}{N}\left\{\left[k-1+\left(\frac{1}{r}-2\right)\right]N_{k-1} - \left[k+\left(\frac{1}{r}-2\right)\right]N_k\right\} + \delta_{k1}. \tag{10.18}$$

Thus uniform attachment, in conjunction with redirection, generates a growing network in which the attachment rate to a node of degree k is a *linear* function of k, albeit with an additive shift. To determine the degree distribution, we now study preferential attachment networks systematically, from which the solution of the redirection model follows easily.

Preferential attachment networks

The degree distribution

Let us again consider networks in which each new node attaches to one pre-existing node of degree k with rate A_k. The master equation for the degree distribution is, in analogy with Eq. (10.15) for the uniformly growing tree,

$$\frac{dN_k}{dN} = \frac{A_{k-1}N_{k-1} - A_k N_k}{A(N)} + \delta_{k1}. \tag{10.19}$$

The first term on the right accounts for processes in which the new node connects to a node that already has $k-1$ links, thereby increasing N_k by one. Since there are N_{k-1} nodes of degree $k-1$, the total rate at which such processes occur equals to $A_{k-1}N_{k-1}$. The factor $A(N) \equiv \sum_{j\geq 1} A_j N_j(N)$ is the total rate for any event to occur, so that $A_{k-1}N_{k-1}/A(N)$ is the probability to attach to a node of degree $k-1$. A corresponding role is played by the second (loss) term on the right-hand side; namely, $A_k N_k/A(N)$ is the probability that the new node connects to a node with k links, thus leading to a loss in N_k by one. The last term accounts for the new node itself that has one outgoing link and no incoming links.

For attachment rates that do not grow faster than linearly with k, both the degree distribution and $A(N)$ grow linearly with time (see problem 10.x). This fact suggests making the substitutions $N_k(N) = N\, n_k$ and $A(N) = \mu N$ in Eq. (10.19). With this step, the overall dependence on N cancels out, leaving behind the recursion relations

$$n_k = \frac{A_{k-1}n_{k-1} - A_k n_k}{\mu} \qquad k > 1, \qquad \text{and} \quad n_1 = -\frac{A_1 n_1}{\mu} + 1, \tag{10.20}$$

with formal solution

$$n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}. \tag{10.21}$$

To make this solution explicit, we need the amplitude μ in $A(N) = \mu N$. Using the definition $\mu = \sum_{j\geq 1} A_j n_j$ in Eq. (10.21), we obtain the condition

$$\sum_{k=1}^{\infty} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1} = 1, \tag{10.22}$$

which shows that the amplitude μ depends on the functional form of the attachment rate. When $A_k = k^\gamma$ with $0 \leq \gamma \leq 1$, a numerical solution of Eq. (10.22) shows that μ varies smoothly between 1 and 2 as γ increases from 0 to 1.

For sublinear attachment rates, $\gamma < 1$, we rewrite the product Eq. (10.21) as the exponential of a sum, convert the sum to an integral, and then expand the logarithm inside the integral in a Taylor series. These steps lead to

$$n_k \sim \begin{cases} k^{-\gamma} \exp\left[-\mu\left(\frac{k^{1-\gamma} - 2^{1-\gamma}}{1-\gamma}\right)\right] & \frac{1}{2} < \gamma < 1, \\ k^{(\mu^2-1)/2} \exp\left[-2\mu\sqrt{k}\right] & \gamma = \frac{1}{2}, \\ k^{-\gamma} \exp\left[-\mu\,\frac{k^{1-\gamma}}{1-\gamma} + \frac{\mu^2}{2}\,\frac{k^{1-2\gamma}}{1-2\gamma}\right] & \frac{1}{3} < \gamma < \frac{1}{2}, \end{cases} \tag{10.23}$$

etc. Whenever γ decreases below $1/m$, with m a positive integer, an additional term in the exponential arises from the now relevant contribution of the next higher-order term in the expansion of the product in Eq. (10.21).

For the strictly linear attachment rate, $A_k = k$, the total event rate is $A = \sum_k A_k N_k = \sum_k k N_k = 2N$. Substituting this value for $A = \mu N$ in Eq. (10.20) and solving the resulting recursion gives the discrete power-law form

$$n_k = \frac{4}{k(k+1)(k+2)} = \frac{4\Gamma(k)}{\Gamma(k+3)} \sim \frac{4}{k^3}. \tag{10.24}$$

The main feature of this result is that there is no natural degree scale. For this reason, such networks have been dubbed *scale free*, and they stand in stark contrast to the delta-function degree distribution of regular lattices and the Poisson degree distribution of the Erdös-Rényi random graph.

A surprising feature of linear preferential attachment is that the exponent of the power-law degree distribution is *non-universal*. The asymptotic evaluation of the product in Eq. (10.21) generally leads to a degree distribution exponent that can assume *any* value greater than 2. All that is required is that the attachment rate is *asymptotically linear*, $A_k \sim k$, rather than strictly linear, $A_k = k$. This non-universal behavior is counter to the conventional wisdom of critical phenomena in which power laws, by their very nature, should not depend on microscopic model details.

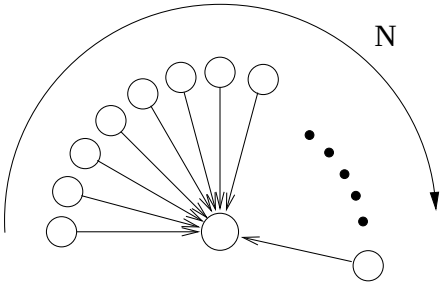


Figure 10.5: Creation of a "bible" for attachment rate $A_k \sim k^\gamma$ with $\gamma > 2$. In the configuration shown each new node attaches only to the bible.

What happens for superlinear attachments rates? Now an analog of gelation occurs in that nearly all the links in the network condense onto a single node, while all other nodes are attached to a small number of links. Especially singular behavior occurs for $\gamma > 2$ because one node there is a non-zero probability that a single node links to *every* node in an infinite network, while only a finite number of links exist between all other nodes. We call such a highly-linked node a "bible". It is easy to see that the probability for a bible to exist is non zero when $\gamma > 2$. Suppose that there is a bible after then network contains $N+1$ nodes (1 bible and N citing nodes). The probability that the next node links to the initial node is then $N^\gamma/(N+N^\gamma)$, and the probability that this connection pattern to continue indefinitely is

$$\mathcal{P} = \prod_{N=1}^{\infty} \frac{1}{1+N^{1-\gamma}} \ .$$

Evaluating this product by the standard steps of writing the product as the exponential of a sum, approximating the sum as an integral, and expanding the logarithm in the integrand to first order, we find that $\mathcal{P} = 0$ for $\gamma \leq 2$ and $\mathcal{P} > 0$ for $\gamma > 2$. Thus for $\gamma > 2$, there a non-zero probability for a bible to exist even in an infinite network.

When $1 < \gamma < 2$, singular behavior still arises in which one node is linked to all but a small number of other nodes. There is a also an infinite sequence of subtle connectivity transitions in the behavior of the number of low-degree nodes. For $3/2 < \gamma < 2$, the number of nodes of degree 2 grows as $N^{2-\gamma}$, while the number of nodes with degree > 2 remains finite. For $4/3 < \gamma < 3/2$, the number of nodes of degree 3 grows as $N^{3-2\gamma}$ and the number with degree > 3 is finite. Generally for $\frac{m+1}{m} < \gamma < \frac{m}{m-1}$, $N_k \sim N^{k-(k-1)\gamma}$ for $k \leq m$, while the number of nodes with degree greater than m links is finite.

Node attractiveness

In many real networked systems, such as the world-wide web, book sales by individuals, scientific publications, *etc.*, not all nodes are equivalent, but rather, some are more attractive than others at their inception. Thus it is natural that the subsequent attachment rate to a node should be a function of both its degree *and* its attractiveness. The master equation approach easily gives the degree distribution for this natural generalization of preferential attachment.

Each node is assigned an initial "attractiveness" $\eta > 0$ that is chosen from a specified distribution $p_0(\eta)$. The attachment rate for a node with degree k and attractiveness η is defined as $A_k(\eta)$. To characterize how nodes evolve, we need to monitor their degree and keep track of their attractiveness. Thus let $N_k(\eta)$ be the number of nodes of degree k and attractiveness η. The evolution of this joint degree-attractiveness distribution is governed by the master equation

$$\frac{dN_k(\eta)}{dN} = \frac{A_{k-1}(\eta)N_{k-1}(\eta) - A_k(\eta)N_k(\eta)}{A} + p_0(\eta)\delta_{k1}, \tag{10.25}$$

where $A = \int d\eta \sum_k A_k(\eta)N_k(\eta)$ is the total rate. Following the same approach as that used to analyze Eq. (10.19), we substitute $A = \mu N$ and $N_k(\eta) = Nn_k(\eta)$ into Eq. (10.25). The solution of the resulting recursion relation is

$$n_k(\eta) = p_0(\eta)\frac{\mu}{A_k(\eta)} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j(\eta)}\right)^{-1}. \tag{10.26}$$

As a simple and generic example, consider the case where $A_k(\eta) = \eta k$, that is, the attachment rate is linear in degree *and* in attractiveness. Applying the same analysis as in the homogeneous network, we obtain the degree distribution

$$n_k(\eta) = \frac{\mu\, p_0(\eta)}{\eta} \frac{\Gamma(k)\,\Gamma\left(1+\frac{\mu}{\eta}\right)}{\Gamma\left(k+1+\frac{\mu}{\eta}\right)}. \tag{10.27}$$

Thus for nodes with a fixed attractiveness η, the asymptotic degree distribution is the power law $n_k(\eta) \sim k^{-1-\mu/\eta}$. What is perhaps more relevant, however, is the degree distribution averaged over the attractiveness distribution. For this purpose, we need the amplitude μ. We therefore substitute (10.27) into the definition $\mu = \int d\eta \sum_{k\geq 1} A_k(\eta)\, n_k(\eta)$ and use the identity

$$\sum_{k=1}^{\infty} \frac{\Gamma(k+u)}{\Gamma(k+v)} = \frac{\Gamma(u+1)}{(v-u-1)\,\Gamma(v)}$$

to yield the implicit relation that determines μ:

$$1 = \int d\eta\, p_0(\eta) \left(\frac{\mu}{\eta} - 1\right)^{-1}. \tag{10.28}$$

The above condition leads to two alternatives: in the pathological case where the support of η is unbounded so that arbitrarily attractive nodes can exist, the integral diverges and there is no solution for μ. In this case, the most attractive node is connected to a finite fraction of all links. Conversely, if the support of η is bounded, then the degree distribution for fixed η is simply the power law $n_k(\eta) \sim k^{-\nu(\eta)}$, with an attractiveness-dependent exponent $\nu(\eta) = 1 + \mu/\eta$. However the degree distribution averaged over all attractiveness, $\langle n_k\rangle = \int d\eta\, n_k(\eta)$, is no longer a power law, but rather $\langle n_k\rangle$ is governed by properties of the initial attractiveness distribution near the upper cutoff. For example, if $p_0(\eta) \sim (\eta_{\max} - \eta)^{\omega-1}$ (with $\omega > 0$ to ensure normalization), the total degree distribution is

$$n_k \sim k^{-(1+\mu/\eta_{\max})}\, (\ln k)^{-\omega}. \tag{10.29}$$

10.3 Finite Networks

Since real networks are necessarily finite, it is wothwhile to ask: what is the role of finiteness on the properties of growing networks? Clearly, finiteness imposes a cutoff on the power-law tail degree distribution (Fig. 10.6), and we wish to quantify this cutoff and related manifestations of finitness. For finite N, the state of a network can be more generally characterized by the set $\mathbf{N} = \{N_1, N_2, \ldots\}$ that specifies the number of nodes N_k of degree k. Each time a new node is introduced into the network, its state $\mathbf{N}$ evolves by:

$$(N_1, N_2) \to (N_1, N_2 + 1),$$
$$(N_1, N_k, N_{k+1}) \to (N_1 + 1, N_k - 1, N_{k+1} + 1).$$

The first process arises when the new node attaches to an existing node of degree 1; in this case, the number of nodes of degree 1 does not change while the number of nodes of degree 2 increases by 1. The second line accounts for the case where the new node attaches to a node of degree $k > 1$. From these processes, it is straightforward, in principle, to write the master equation for the joint probability distribution $P(\mathbf{N})$. However, such an equation would provide much more information than is of practical interest. Here we focus on the degree distribution and fluctuations in the degree distribution. As we shall show, for a finite network the number of nodes of fixed degree, $N_k(N)$, are random variables that become sharply peaked about their average values in the $N \to \infty$ limit.

When the number of nodes N is finite, there will also necessarily be a maximal degree $k_{\max}$. Thus predictions for the degree distribution that are implicitly based on an infinite network must eventually break down as k approaches $k_{\max}$. The maximal degree may be determined by the extremal criterion $\sum_{k\geq k_{\max}} \langle N_k(N)\rangle \approx 1$ which states that there should be one node whose degree is in the range $(k_{\max}, \infty)$. This criterion yields $k_{\max} \sim N^{1/(\nu-1)}$ when the degree distribution of an infinite network asymptotically decays as $k^{-\nu}$.

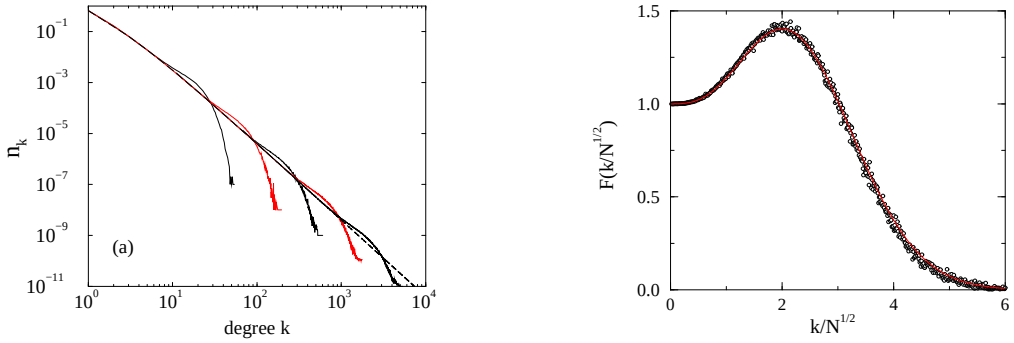


Figure 10.6: (left) Normalized degree distributions for strictly linear preferential attachment networks ($A_k = k$) with $10^2, 10^3, \ldots, 10^6$ links (upper left to lower right). The dashed line is the asymptotic result $n_k = 4/[k(k+1)(k+2)]$. (right) The scaling function $F(\xi)$, with $\xi = k/N^{1/2}$ from Eq. (10.39). The circles give the simulation data of 10^6 realizations of a network with $N = 10^4$ links for the dimer initial condition.

Nodes of Fixed Degree

Degree 1

To appreciate the role of finiteness in the simplest possible setting, consider the number of nodes of degree 1, $N_1(N)$. For the case of strictly linear preferential attachment, $A_k = k$, the average number of such nodes evolves according to

$$\langle N_1(N+1)\rangle = \left\langle N_1(N) \times \frac{N_1(N)}{2N} \right\rangle + \left\langle (N_1(N)+1) \times \left(1 - \frac{N_1}{2N}\right) \right\rangle .$$

The first term on the right accounts for the new node attaching to a node of degree one, an event that occurs with probability $N_1/2N$. In this case, the number of nodes of degree one does not change. The second term accounts for the new node attaching to a node of degree greater than one with probability $(1 - N_1/2N)$. For this event N_1 increases by one. Simplifying, we obtain the recursion

$$\langle N_1(N+1)\rangle = 1 + \left(1 - \frac{1}{2N}\right) \langle N_1(N)\rangle. \tag{10.30}$$

We take the initial condition $\langle N_1(1)\rangle = N_1(1) = 2$. **why?**

To solve this recursion, we multiply (10.30) by Nw^{N-1} and sum over $N \geq 1$ to convert it into the differential equation for the generating function $\mathcal{X}_1(w) = \sum_{N\geq 1}\langle N_1(N)\rangle\, w^{N-1}$,

$$\frac{d\mathcal{X}_1}{dw} = \frac{1}{(1-w)^2} + \frac{1}{2}\,\mathcal{X}_1 + w\,\frac{d\mathcal{X}_1}{dw}\,. \tag{10.31}$$

Solving Eq. (10.31) subject to the initial condition $\mathcal{X}_1(0) = 2$ gives $\mathcal{X}_1(w) = \frac{2}{3}(1-w)^{-2} + \frac{4}{3}(1-w)^{-1/2}$, and expanding this solution in a Taylor series in w leads to

$$\langle N_1(N)\rangle = \frac{2}{3}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)}\,. \tag{10.32}$$

Thus the fraction of nodes of degree 1 approaches the expected value of 2/3 [see Eq. (10.24)], but with corrections that vanish as $N^{-1/2}$.

Degree Greater Than One

Following the same reasoning as in the case of nodes of degree 1, the number of nodes of degree $k \geq 2$, $N_k(N)$ evolves as

$$\langle N_k(N+1)\rangle = \left\langle N_1(N)\times\frac{N_1(N)}{2N}\right\rangle + \left\langle (N_1(N)+1)\times\left(1-\frac{N_1}{2N}\right)\right\rangle\,.$$

The first term on the right accounts for the new node attaching to a node of degree k with probability $N_k/2N$, after which the N_k decreases by 1. The second term accounts for the new node attaching to a node of degree $k-1$ with probability $N_{k-1}/2N$), after which N_k increases by 1. The last term accounts for attachment to all other nodes, which leads to no change in N_k. This then gives the recursion

$$\langle N_k(N+1)\rangle = \langle N_k(N)\rangle + \left\langle\frac{(k-1)N_{k-1}(N) - kN_k(N)}{2N}\right\rangle. \tag{10.33}$$

The first term on the right accounts for the new node attaching to a node of degree $k-1$ with probability $N_{k-1}/2N$. The second term accounts for the new node attaching to a node of degree k with probability $N_k/2N$. For this event N_1 increases by one.

Thus recursion can be again solved by the generating function method (problem 10.x). Expanding this generating function in a Taylor series we then obtain $\langle N_k(N)\rangle$, and the explicit results for the first few k are:

$$\begin{aligned}
\langle N_1(N)\rangle &= \frac{2}{3}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)},\\
\langle N_2(N)\rangle &= \frac{1}{6}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)} - \frac{3}{2}\,\delta_{N,1},\\
\langle N_3(N)\rangle &= \frac{1}{15}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)} - \frac{4}{5\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{3}{2}\right)}{\Gamma(N)} - 3\,\delta_{N,1}
\end{aligned}$$

Thus as N increases the number of nodes N_k with fixed degree k approaches the value obtained for the infinite system, $\frac{4}{k(k+1)(k+2)}$, but with corrections whose leading behavior is proportional to $N^{-1/2}$.

Nodes of Arbitrary Degree

More generally, what is the dependence of N_k on *both* k and on N? The existence of this maximal degree suggests that the degree distribution should be described by the finite-size scaling form

$$\langle N_k(N)\rangle \simeq N n_k F(\xi), \qquad \xi = k/k_{\max}. \tag{10.34}$$

This scaling function has a well-defined peak for $\xi \approx 2$ (Fig. 10.6), and through the use of the generating function method, it is possible to determine the full behavior of the scaling function. To start, we need the dependence of $N_k(N)$ on both N nd k. It is therefore useful to introduce the two-variable generating function

$$\mathcal{N}(w,z) = \sum_{N=1}^{\infty}\sum_{k=1}^{\infty}\langle N_k(N)\rangle\, w^{N-1}\, z^k\,. \tag{10.35}$$

Taking Eq. (10.33), multiplying by $w^{N-1}z^k$ and summing over all N and k, the generating function $\mathcal{N}(w,z)$ satisfies

$$\left(2(1-w)\,\frac{\partial}{\partial w} + z(1-z)\,\frac{\partial}{\partial z} - 2\right)\mathcal{N} = \frac{2z}{(1-w)^2}\;. \tag{10.36}$$

To simplify this equation, we introduce the rotated variables x, y defined by $x+y = -\frac{1}{2}\ln(1-w)$ and $x-y = \ln\frac{z}{1-z}$, to recast Eq. (10.36) into

$$\left(\frac{\partial}{\partial x} - 2\right)\mathcal{N}(x,y) = \frac{2\,e^{5x+4y}}{e^x+e^y}\;, \tag{10.37}$$

whose general solution is

$$\mathcal{N}(x,y) = e^{4x+4y} - 2e^{3x+5y} - e^{2x+2y} + 2e^{2x+4y} + 2\,\frac{e^{2x+2y}}{1+e^{2y}} + 2e^{2x+6y}\ln\left(\frac{e^{x+y}+e^{2y}}{1+e^{2y}}\right). \tag{10.38}$$

We may then extract the scaling function $F(\xi)$ from this generating function, and the final result is

$$F(\xi) = \operatorname{erfc}\left(\frac{\xi}{2}\right) + \frac{2\xi+\xi^3}{\sqrt{4\pi}}\;e^{-\xi^2/4}\,, \tag{10.39}$$

where $\operatorname{erfc}(x)$ is the complementary error function. The most important feature of this result is that the exact average degree distribution has a Gaussian large-degree tail

$$\langle N_k(N)\rangle \to \frac{2}{\sqrt{\pi N}}\;e^{-k^2/4N}\,, \tag{10.40}$$

and moreover, the scaling function in Eq. (10.39) completely describes the finite-size correction to the degree distribution (Fig. 10.6).

Higher Moments and Their Fluctuation

We now turn to higher moments of the degree distribution. While the zeroth and first moments of the degree distribution are simply related to the total number of links for *any* network topology, the higher moments are not so simply characterized, but instead reflect the power-law tail of the degree distribution.

Using the exact expression (10.38) for the generating function, the second moment of the degree distribution is obtained from

$$\left(z^2\,\frac{\partial}{\partial z}\right)^2 \mathcal{N}(w,z)\Big|_{z=1} = \;\frac{4-2\ln(1-w)}{(1-w)^2}\;.$$

We now expand the right-hand side in a series in w to yield

$$\langle k^2\rangle \equiv \sum_{k=1}^{\infty} k^2\langle N_k\rangle = 2NH_N, = 2N\ln N + 2\gamma N + 1 - \frac{1}{6N} + \ldots\,, \tag{10.41}$$

where $H_N = \sum_{1\le j\le N} j^{-1}$ is the harmonic number and $\gamma \approx 0.5772166$ is Euler's constant.

For the third moment we find

$$\langle k^3\rangle = \frac{32}{\sqrt{\pi}}\,\frac{\Gamma\left(N+\frac{3}{2}\right)}{\Gamma(N)} - 6NH_N - 16N\,. \tag{10.42}$$

More generally, the dependence of the moments on N stems from the power-law tail of the degree distribution $\langle N_k\rangle \propto N/k^3$. From this asymptotic distribution, a suitably normalized set of measures for the mean degree

$$\mathcal{M}_n = \left(\frac{\langle k^n\rangle}{\langle k^0\rangle}\right)^{1/n}, \tag{10.43}$$

has the following N dependence:

$$\mathcal{M}_n \propto \begin{cases} \text{const.} & n<2 \\ \ln N & n=2 \\ N^{(n-2)/2} & n>2 \end{cases} \tag{10.44}$$

Sensitivity

To illustrate the crucial role of the initial condition, let's study the degree of the first node in the network. Let $P(k,N)$ be the probability that the first node has degree k in a network of N nodes. For linear preferential attachment, $A_k = k$, this probability obeys the master equation

$$P(k,N+1) = \frac{k-1}{2N}\,P(k-1,N) + \frac{2N-k}{2N}\;P(k,N). \tag{10.45}$$

The first term on the right accounts for the situation when the first node has degree $k-1$: a new node can attach to it with probability $(k-1)/2N$, thereby increasing the probability for the first node to have degree k. Conversely, with probability $(2N-k)/2N$ a new node does not attach to the earliest node, thereby giving the second term on the right.

The solution to the master equation (10.45) for the "dimer" initial condition ●—● is

$$P(k,N) = \frac{1}{2^{2N-k-1}}\,\frac{(2N-k-1)!}{(N-k)!\,(N-1)!} \longrightarrow \frac{1}{\sqrt{\pi N}}\;e^{-k^2/4N}, \tag{10.46}$$

where the asymptotic behavior applies in the limit $N\to\infty$, with the scaling variable $k/N^{1/2}$ being finite. Thus the average degree of the first node is $\langle k\rangle_1 = \sqrt{4N/\pi} \approx 1.228\sqrt{N}$. On the other hand, from the extremal criterion for the largest degree in the network

$$\sum_{k_{\max}}^{\infty} N n_k = 1,$$

and using asymptotic degree distribution $n_k \sim 4/k^3$, we obtain the largest degree $k_{\max}\sim\sqrt{2N}\approx 1.4142\sqrt{N}$. Thus the degree of the first node of the network is close the largest degree; this fact implies that there is a substantial probability that the first node in the network is the one with the largest degree.

Although $P(k,N)$ contains all information about the degree of the first node, its moments $\langle k^a\rangle_N = \sum k^a P(k,N)$ are simpler to appreciate. Using Eq. (10.45), the average degree of the initial node satisfies the recursion relation

$$\langle k\rangle_{N+1} = \langle k\rangle_N\left(1+\frac{1}{2N}\right), \tag{10.47}$$

whose solution is

$$\langle k\rangle_N = \Lambda\,\frac{\Gamma\left(N+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(N)} \sim \frac{\Lambda}{\sqrt{\pi}}\,N^{1/2}. \tag{10.48}$$

The prefactor Λ depends on the initial condition, with $\Lambda = 2, 8/3, 16/5, \ldots$ for the dimer, trimer, tetramer, *etc.*, initial conditions.

This multiplicative dependence on the initial condition means that the first few growth steps substantially affect the average degree of the first node. For example, for the dimer initial condition, the average degree of the first node is, asymptotically, $\langle k\rangle_N \sim 2\sqrt{N/\pi}$. However, if the second link attaches to the first node, an effective trimer initial condition arises and $\langle k\rangle_N \sim (8/3)\sqrt{N/\pi}$. Thus small initial perturbations at the beginning of the network growth lead to huge differences in the degree of the first node.

Problems

Section 10.1

1. Solve the master equations for the degree distribution of the ER graph

$$\frac{dn_k}{dt} = n_{k-1} - n_k$$

one by one for the initial condition $n_0(t=0) = 1$ and show that the solution is

$$n_k = \frac{t^k}{k!}\, e^{-kt}.$$

Section 10.2

1. Compute the degree distribution for the uniformly growing tree for a few values of $N = 1, 2, 3, \ldots$ within the (i) continuous and (ii) discrete approach. For the continous formulation, solve the system of differential equations (10.15). For the discrete approach, solve the exact recursion formulae

$$N_k(N+1) - N_k(N) = \frac{N_{k-1}(N) - N_k(N)}{N} + \delta_{k,1}.$$

Compare your results with the asymptotic average degree distribution $N_k(N) \sim N2^{-k}$.

2. Show that the zeroth and first moments of the degree distribution $M_n(N) \equiv \sum_{j\geq 1} j^n N_j(N)$ have the time dependence $M_0(N) = N$ and $M_1(N) = 2N$, independent of the attachment rate A_k.

3. Verify, for attachment rates that do not grow faster than linearly with k, that both the degree distribution $N_k(N)$ and $A(N)$ both grow linearly with time.

4. Show that the limiting behavior of μ in $A(N) = \mu N$ is give by:

$$\begin{aligned} \mu &= 1 + B_0\gamma + \mathcal{O}\left(\gamma^2\right), \qquad \gamma \downarrow 0 \\ \mu &= 2 - B_1(1-\gamma) + \mathcal{O}\left((1-\gamma)^2\right), \qquad \gamma \uparrow 1 \end{aligned}$$

with

$$\begin{aligned} B_0 &= \sum_{j=1}^{\infty} \frac{\ln j}{2^j} = 0.5078\ldots, \\ B_1 &= 4\sum_{j=1}^{\infty} \frac{\ln j}{(j+1)(j+2)} = 2.407\ldots. \end{aligned}$$

Here γ is the exponent in the attachment rate A_k defined by $A_k = k^\gamma$.

5. Determine the degree distribution for the shifted linear attachment rate $A_k = k + \lambda$. First show that $A(N) = \sum_j A_j N_j(N)$ now equals $A(N) = M_1(N) + \lambda M_0(N)$. Using these results in the master equation show that the degree distribution is

$$n_k = (2+\lambda)\,\frac{\Gamma(3+2\lambda)}{\Gamma(1+\lambda)}\,\frac{\Gamma(k+\lambda)}{\Gamma(k+3+2\lambda)}. \tag{10.49}$$

Show that asymptotically, this distribution decays as $k^{-\nu}$, with $\nu = 3 + \lambda = 1 + \frac{1}{r}$.

6. Consider the connection kernel $A_1 = 1$ and $A_k = ak$ for $k \geq 2$. Show that the resulting degree distribution is asymptotically a power law, $n_k \sim k^{-\nu}$, with $\nu = (3 + \sqrt{1 + 8/a_\infty})/2$, which can indeed be tuned to any value larger than 2.

7. Generalize linear preferential attachment networks to the case where each new node links to m pre-existing nodes. Write the master equation for this process, and by applying the same approach as that used for Eq. (10.19), find the the degree distribution

8. Solve the recursion Eq. (10.33) for the number of nodes of degree k, $N_k(N)$ by the generating function method.

   Solution: Define the generating function as $\mathcal{X}_k(w) = \sum_{N=1}^{\infty} \langle N_k(N) \rangle \, w^{N-1}$. We multiple Eq. (10.33) by w^N to convert it to a differential equation that relates $\mathcal{X}_k$ and $\mathcal{X}_{k-1}$. This equation is further simplified by making the transformation

$$\mathcal{X}_k(w) = (1-w)^{\frac{k}{2}-1} \, \mathcal{U}_k(u), \quad u = \frac{1}{\sqrt{1-w}} - 1. \tag{10.50}$$

   The resulting equation is

$$\frac{d\mathcal{U}_k}{du} = (k-1)\, \mathcal{U}_{k-1}, \qquad k \geq 2. \tag{10.51}$$

   Rewriting our previous solution for $\mathcal{X}_1$ as

$$\mathcal{U}_1(u) = \frac{2}{3}\, u^3 + 2u^2 + 2u + 2, \tag{10.52}$$

   one can solve Eqs. (10.51) subject to the initial condition $\mathcal{U}_k(u=0) = 0$ for $k \geq 2$. The final result is

$$\mathcal{U}_k(u) = \frac{4u^{k+2}}{k(k+1)(k+2)} + \frac{4u^{k+1}}{k(k+1)} + \frac{2u^k}{k} + 2u^{k-1}\,.$$

   Using the binomial formula, we transform $\mathcal{X}_k(z)$ into the series

$$\begin{aligned} \mathcal{X}_k(w) = {} & \frac{4}{k(k+1)(k+2)} \, \frac{1}{(1-w)^2} + \frac{4}{3} \, \frac{1}{(1-w)^{1/2}} \\ & + 2 \sum_{a=1}^{k-1} (-1)^a \, \frac{a+2}{a+3} \binom{k-1}{a} (1-w)^{(a-1)/2}\,. \end{aligned}$$

Appendix A

MATTERS OF TECHNIQUE

A.1 Transform Methods

Laplace transforms for continuum systems

Generating function for discrete systems

This method is demonstrated for the Fibonacci sequence

$$F_n = F_{n-1} + F_{n-2} \tag{A.1}$$

for $n \geq 1$ with $F_{-1} = 0$ and $F_0 = 1$. The first six terms in the sequence are $\{1, 1, 2, 3, 5, 8\}$. The generating function is defined as follows

$$F(z) = \sum_{n=0}^{\infty} F_n z^n. \tag{A.2}$$

To obtain the generating function, we multiply the recursion relation by z^n and sum over n. The left hand side yields $F(z) - 1$. The two terms on the right hand side yield $zF(z)$ and $z^2F(z)$, respectively. Thus, the generating function satisfies $F(z) = 1 + (z + z^2)F(z)$ and its solution is immediate

$$F(z) = \frac{1}{1 - z - z^2}. \tag{A.3}$$

We can manually check that this expression reproduces the first few sequence elements, $F(z) = 1 + z + 2z^2 + 3z^3 + 5z^4 + \cdots$. The generating function is re-written as follows

$$F(z) = \frac{1}{1 - z - z^2} = \frac{1}{(1 - \lambda_+ z)(1 - \lambda_- z)} = \frac{1}{\lambda_+ - \lambda_-} \left[\frac{\lambda_+}{1 - \lambda_+ z} - \frac{\lambda_-}{1 - \lambda_- z} \right]. \tag{A.4}$$

Here, $\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}$. Expanding the two fractions as Taylor series, the Fibonacci numbers follow

$$F_n = \frac{\lambda_+^{n+1} - \lambda_-^{n+1}}{\lambda_+ - \lambda_-}. \tag{A.5}$$

A.2 Relation between Laplace Transforms and Real Time Quantities

In many time-dependent phenomena, we want the asymptotic behavior of some quantity as a function of time when only its generating function or its Laplace transform is available. A typical example is a function $F(t)$ whose Laplace transform has the small-s behavior $F(s) \sim s^{\mu-1}$, or equivalently, whose generating function has the form $F(z) \sim (1 - z)^{\mu-1}$, with $\mu < 1$ as $z \to 1$ from below. We will show that the corresponding

time dependence is $F(t) \sim t^{-\mu}$ as $t \to \infty$. The first point to make is that the generating function and the Laplace transform are equivalent. For a function $F(t)$ that is defined only for positive integer values, that is $t = n$, with n an integer, the Laplace transform is

$$F(s) = \int_0^\infty F(t)\, e^{-st}\, dt = \sum_{t=0}^{\infty} F(n)\, e^{-sn}.$$

Thus defining $z = e^{-s}$, the Laplace transform is the same as the generating function. In the limit $z \to 1$ from below, the sum cuts of only very slowly and can be replaced by an integral. This step leads to the Laplace transform in the limit $s \to 0$.

If we are interested only in long-time properties of a function $F(t)$, these features can be obtained through simple and intuitively appealing means. While lacking in rigor, this approach provides the correct behavior for all cases of physical interest. The most useful of these methods are outlined in this section. Let us first determine the long-time behavior of a function $F(t)$ when only its Laplace transform is known. There are two fundamentally different cases to consider: (a) $\int_0^\infty F(t)\, dt$ diverges or (b) $\int_0^\infty F(t)\, dt$ converges.

In the former case, relate the Laplace transform $F(s)$ to $F(t)$ by the following simple step:

$$F(s) = \int_0^\infty F(s)\, e^{-st}\, dt \approx \int^{t^*} F(t)\, dt. \tag{A.6}$$

That is, we simply replace the exponential cutoff in the integral, with characteristic lifetime $t^* = 1/s$, by a step function at t^*. Although this crude approximation introduces numerical errors of the order of 1, the essential asymptotic behavior of $F(t)$ is preserved when $\int 0^\infty F(t)\, dt$ diverges. Now if $F(t) \to t^{-\mu}$ with $\mu < 1$ as $t \to \infty$, then $F(s)$ diverges as

$$F(s) \sim \int^{1/s} t^{-\mu}\, dt \sim s^{\mu-1} \tag{A.7}$$

as $s \to 0$ from below. In summary, the fundamental connection when $\int_0^\infty F(t)\, dt$ diverges is

$$F(t) \sim t^{-\mu} \quad \longleftrightarrow \quad F(s) \sim s^{\mu-1}. \tag{A.8}$$

The above result also provides a general connection between the time integral of a function, $\mathcal{F}(t) \equiv \int_0^t F(t)\, dt$, and the Laplace transform of F. For $s = 1/t^*$ with $t^* \to \infty$, Eq. (A.7) is just the following statement:

$$F(s = 1/t^*) \sim \int_0^{t^*} F(t)\, dt = \mathcal{F}(t^*). \tag{A.9}$$

Thus a mere variable substitution provides an approximate, but asymptotically correct, algebraic relation between the Laplace transform of a function and the time integral of this same function. For this class of examples, there is no need to perform an integral to relate a function and its Laplace transform.

In the opposite situation where $\int_0^\infty F(t)\, dt = \mathcal{F}(\infty)$ converges, we can obtain the connection between $F(t)$ and $F(s)$ in a slightly different way. Let us again suppose that $F(t) \sim t^{-\mu}$ as $t \to \infty$, but now with $\mu > 1$ so that $F(s)$ is finite as $s \to 0$. Exploiting the fact that the time integral of $F(t)$ converges, we write

$$\begin{aligned} F(s) &= \int^\infty t^{-\mu} \left[1 - (1 - e^{-st})\right]\, dt, \\ &\sim \mathcal{F}(\infty) + \int_{1/s}^\infty t^{-\mu}\, dt, \\ &\sim \mathcal{F}(\infty) + s^{\mu-1}. \end{aligned} \tag{A.10}$$

Again, we replace the exponential cutoff in the integrand by a sharp cutoff.

In summary, the small-s behavior of the Laplace transform, or, equivalently, the $z \to 1$ behavior of the generating function, are sufficient to determine the long-time behavior of the function itself. Since the transformed quantities are usually easy to obtain by the solution of an appropriate boundary-value problem, the asymptotic methods outlined here provide a simple route to obtain long-time behavior.

In the context of time-dependent phenomena, one of the most useful features of the Laplace transform is that it encodes all positive integer powers of the mean time. That is, we define the positive integer moments of $F(t)$ as

$$\langle t^n \rangle = \frac{\int_0^\infty t^n \, F(t) \, dt}{\int_0^\infty F(t) \, dt}. \tag{A.11}$$

If all these moments exist, then $F(s)$ can be written as a Taylor series in s These generate the positive integer moments of $F(t)$ by means of

$$\begin{aligned} F(s) &= \int_0^\infty F(t) \, e^{-st} \, dt \\ &= \int_0^\infty F(t) \, \left(1 - st + \frac{s^2 t^2}{2!} - \frac{s^3 t^3}{3!} + \dots \right) \\ &= \mathcal{F}(\infty) \left(1 - s\langle t \rangle + \frac{s^2}{2!} \langle t^2 \rangle - \frac{s^3}{3!} \langle t^3 \rangle + \dots \right). \end{aligned} \tag{A.12}$$

Thus the Laplace transform is a *moment generating function*, as it contains *all* the positive integer moments of the probability distribution $F(t)$.

A.3 Asymptotic Analysis

This method is demonstrated for the Fibonacci numbers, defined according to the recursion rule (A.1). How do these numbers grow with n? The simpler series $F_n = 2F_{n-1}$ with $F_0 = 1$ is simply the geometric series $F_n = 2^n$. Thus, we try

$$F_n \sim \lambda^n \tag{A.13}$$

and expect $z < 2$. Substituting this ansatz into the recurrence relation (A.1) shows that the ansatz is compatible with the equation when $\lambda^2 = \lambda + 1$. There are two roots $\lambda_\pm = \left(1 \pm \sqrt{5}\right)/2$. The largest root is the relevant one asymptotically, $\lambda = \lambda_+$. This number, the golden ratio, characterizes the growth of the sequence

$$\lambda = \frac{1+\sqrt{5}}{2}. \tag{A.14}$$

Thus, the growth of the sequence is obtained up to the proportionality coefficient $\lim_{n\to\infty} F_n/\lambda^n$. This prefactor is of secondary importance, and it is of the order unity. In this case, an exact solution of F_n was possible, but in typical applications, exact solutions are cumbersome and asymptotic analysis are quite useful.

Steepest Descent

A.4 Scaling Approaches

Separation of variables

Conversion of PDEs to ODEs

A.5 Differential Equations

We demonstrate several useful methods for the same problem: aggregation with constant kernel. The Smoluchowsky equation) for c_k, the density of clusters of size k is

$$\frac{dc_k}{dt} = \sum_{i+j=k} c_i c_j - 2N c_k \tag{A.15}$$

and the initial conditions $c_k(0) = \delta_{k,0}$.

Moments methods

The moments are defined as follows

$$M_n(t) = \sum_k x^n c_k. \tag{A.16}$$

Multiplying the rate equations by k^n and summing over k, the moments obey the closed set of equations

$$\frac{dM_n}{dt} = \sum_{l=0}^{n} \binom{n}{l} M_l M_{n-l} - 2M_0 M_n. \tag{A.17}$$

Here, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ are the binomial coefficients. The zeroth moment satisfies $\frac{dM_0}{dt} = -M_0^2$ and the initial conditions $M_0 = 1$, so the solution is $M_0 = (1+t)^{-1}$. The first moment (total mass) is conserved, $M_1 = 1$.

The rest of the moments can be solved recursively. However, such an approach may not be practical. Instead, we solve for the leading behavior of the moments in the long time limit. We assume that average mass, $\langle k\rangle = M_1/M_0$, characterizes all moments, $M_n \sim M_0(M_1/M_0)^n$, with the normalization assuring that the zeroth order moment is recovered. Since asymptotically $\langle k\rangle \simeq t$, the scaling ansatz is

$$M_n \simeq A_n\, t^{n-1} \tag{A.18}$$

with $A_0 = A_1 = 1$. Substituting this ansatz into the moment equation, these coefficients satisfy the recursion relation

$$(n+1)\frac{A_n}{n!} = \sum_{l+m=n} \frac{A_l}{l!}\frac{A_m}{m!}. \tag{A.19}$$

To solve this equation we introduce the generating function

$$A(z) = \sum_n \frac{A_n}{n!} z^n. \tag{A.20}$$

Given the structure of the recursion relation, we conveniently absorbed the factor $n!$ into the definition of the generating function. Multiplying the recursion relation by z^n and summing over n, the generating function satisfies

$$\frac{d}{dz}\left[zA(z)\right] = A^2(z). \tag{A.21}$$

and the boundary condition $A(0) = 0$. This equation can be re-written as the Ricatti equation $zA_z + A = A^2$. Thus, we introduce $g(z) = A^{-1}(z)$ which satisfies $zg_z - g = -1$. The solution to this equation subject to $g(0) = 1$ is $g(z) = 1 - z$. Therefore,

$$A(z) = \frac{1}{1-z}. \tag{A.22}$$

This yields the coefficients $A_n = n!$ and therefore, the leading asymptotic behavior of the moments

$$M_n \simeq n! t^{n-1}. \tag{A.23}$$

The asymptotic form of the size distribution can be obtained from the asymptotic form of the moments. Since the scale $k \sim t$ characterizes the moments, the size distribution attains the form $c_k \sim t^{-2}\Phi(kt^{-1})$. This form is consistent with the moment behavior with the corresponding coefficients merely the moments of the scaling distribution $A_n = \int dx x^n \Phi(x)$. In this case, the coefficients are of a simple form and thus allow us to guess the distribution: $n! = \int dx x^n \exp(-x)$ so $\Phi(x) = \exp(-x)$.

Scaling Solutions

Usually it is much simpler to solve for the scaling function rather than seek a complete exact solution. The exact moment solutions $M_0 = (1+t)^{-1}$ and $M_1 = 1$ show that the average size grows linearly with time $\langle k\rangle = 1 + t \simeq t$. The scaling ansatz

$$c_k \simeq t^{-2}\Phi(kt^{-1}) \tag{A.24}$$

is consistent with the first two moments when $\int dz\, \Phi(z) = \int dz\, z\, \Phi(z) = 1$. Replacing the sum in the rate equation by an integral and substituting this scaling form leads to the following integro-differential equation

$$-z\frac{d}{dz}\Phi(z) = \int_0^z dy\, \Phi(x)\Phi(z-x). \tag{A.25}$$

The convolution structure suggests the Laplace transform: $\hat\Phi(s) = \int dz\, e^{-sz}\Phi(z)$ with the normalization dictating $\hat\Phi(s)|_{s=0} = 1$ and $\frac{d}{ds}\hat\Phi(s)|_{s=0} = -1$. The Laplace transform obeys

$$\frac{d}{ds}\left[s\,\hat\Phi(s)\right] = \hat\Phi^2(s). \tag{A.26}$$

This Ricatti equation is linearize via the transformation $g = 1/\hat\Phi$: $-s\frac{d}{ds}g + g = 1$. With the boundary conditions $g(s)|_{s=0} = \frac{d}{ds}g(s)|_{s=0} = 1$ the solution is $g(s) = 1 + s$ and therefore, $\hat\Phi(s) = (1+s)^{-1}$. The inverse Laplace transform is the exponential

$$\Phi(x) = \exp(-x). \tag{A.27}$$

A.6 Partial Differential Equation

Traveling Waves & Extremum Selection

In many situations, a solution to a partial differential equation attains a traveling wave form asymptotically, thereby reducing the problem complexity. We describe a handy technique for obtaining the propagation velocity of the traveling wave.

Consider the classic Fisher-Kolmogorov-Petrovsky-Piskunov (FKPP) population equation

$$u_t = u_{xx} + u(1-u). \tag{A.28}$$

This partial differential equation describes changes in the average population $u(x,t)$ at position x at time t due to diffusion, birth (linear term), and death as a result of overpopulation (non-linear term). Furthermore, consider the step initial condition $u_0(x) = \Theta(x)$ with $\Theta(x)$ the Heaviside step function: $\Theta(x) = 1$ for $x < 0$ and $\Theta(x) = 0$ for $x > 0$.

There are two steady state solutions: an unstable solution, $u = 0$, and a stable solution, $u = 1$. One anticipates that the stable phase penetrates the unstable one. Furthermore, assume that asymptotically (as $t \to \infty$) the stable phase propagates into the unstable one with a constant velocity c

$$u(x,t) \to F(x - ct). \tag{A.29}$$

The wave-function $F(z)$ satisfies the ordinary differential equation

$$F_{zz} + c\,F_z + F(1-F) = 0 \tag{A.30}$$

and the boundary conditions $F(-\infty) = 1$ and $F(\infty) = 0$. The wave front is located at $x \approx c\,t$. Far ahead of this position, the population is very sparse, so the nonlinear term is negligible. The resulting linear equation $F_{zz} + cF_z + F = 0$ implies an exponential front, $F(z) \propto \exp(-\nu z)$ with the decay coefficient ν satisfying $\nu^2 - c\nu + 1 = 0$. Alternatively, the propagation velocity and the decay coefficient are related via

$$c = \frac{\nu^2 + 1}{\nu}. \tag{A.31}$$

This curve has a minimum at $c_{\min} = 2$, realized when $\nu = \nu_{\min} = 1$. Even though there is a continuous family of solutions, characterized by the velocity c, the minimal value

$$c = 2 \tag{A.32}$$

is actually selected by the dynamics[1]!

This "extremum selection principle" is ubiquitous and very handy. It yields the propagation velocity and reduces the partial differential equation (A.28) into the ordinary differential equation (A.30). This is a considerable simplification from a dependence on two variable x and t to a dependence on only one $z = x - ct$, a transformation equivalent to a scaling transformation. In its core, though it is a *linear* analysis method.

Special solutions in the family of solutions can be still realized. When the initial condition is not compact, but rather has the extended tail, $u_0(x) \sim \exp(-\nu z)$, then the velocity is given by (A.31). This selection is sometimes termed "weak selection".

A.7 Extreme Statistics

Extremes of simple distributions

Lifshitz tail argument

Combining an exact solution of a sub-class of extreme events with the scaling behavior results in a powerful heuristic tool for characterizing extremal statistics.

[1] This can be justified rigorously using a saddle point analysis in the complex plane of a wave dispersion relation equivalent to (A.31).

Consider the discrete time random walk. At time t, the walker is at position $x(t)$, defined recursively via

$$x(t+1) = \begin{cases} x(t) - 1 & \text{with probability } 1/2; \\ x(t) + 1 & \text{with probability } 1/2. \end{cases} \tag{A.33}$$

The walker starts at the origin $x(0) = 0$. The average displacement satisfies the recursion $\langle x(t+1)\rangle = \langle x(t)\rangle + \frac{1}{2} \times (-1) + \frac{1}{2} \times (1) = \langle x(t)\rangle$ and therefore, the average vanishes $\langle x(t)\rangle = 0$. The variance in the displacement satisfies the recursion $\langle x^2(t+1)\rangle = \frac{1}{2}\langle (x(t)-1)^2\rangle + \frac{1}{2}\langle (x(t)+1)^2\rangle = \langle x^2(t)\rangle + 1$. Therefore, the variance grows linearly in time $\langle x^2(t)\rangle = t$, a diffusive behavior. Given the growth $x \simeq \sqrt{t}$ we assume that $P(x,t)$, the probability that the walker is at position x at time t, becomes self-similar

$$P(x,t) \simeq \frac{1}{\sqrt{t}} \Phi\left(\frac{x}{\sqrt{t}}\right). \tag{A.34}$$

An easy to analyze sub-class of extreme random walkers are those that always take a positive step. These walkers move ballistically, $x = t$, and their likelihood is $P(x = t, t) = 2^{-t}$. This probability decreases exponentially with time

$$P(x = t, t) \sim \exp(-Ct). \tag{A.35}$$

Matching the scaling behavior (A.34) with the tail (A.35) shows that the scaling function satisfies $\Phi(\sqrt{t}) \sim \exp(-C\,t)$ (Here, we ignored the secondary algebraic correction $1/\sqrt{t}$.). This yields the tail behavior

$$\Phi(z) \sim \exp(-Cz^2) \tag{A.36}$$

as $z \to \infty$. Thus, we have obtained the functional form of the tail. (This approach does not produce the proportionality constant C.) The moral is that seemingly trivial and partial description, in this case, ballistic walkers, can yield via proper manipulation useful generic information.

Reversion of Series

A.8 Probability theory

Generating functions

Moments and cumulants

Elementary distributions: Normal & Poisson

References

1. M. Abramowitz and A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
2. C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, Singapore, 1984).
3. C. Godrèche and P. Manneville, *Hydrodynamics and Nonlinear Instabilities* (Cambridge University Press, Cambridge, 1998).
4. J. D. Murray, *Mathematical Biology* (Springer-Verlag, Berlin, 1989).
5. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science* (Reading, Mass.: Addison-Wesley, 1989).
6. H. S. Wilf, *generatingfunctionology* (Academic Press, Boston, 1990).

Appendix B

Formulas & Distributions

B.1 Useful Formulas

- The Gamma function

$$\Gamma(n) = \int_0^\infty dx\, x^{n-1} e^{-x}. \tag{B.1}$$

  1. Large-n asymptotics (the Stirling formula)

$$\Gamma(n) \simeq \sqrt{2\pi n}\, n^n e^{-n}. \tag{B.2}$$

  2. Recursion relation

$$\Gamma(n+1) = n\Gamma(n) \tag{B.3}$$

  3. Asymptotic ratio

$$\Gamma(x+a)/\Gamma(x) \simeq x^a \tag{B.4}$$

- The incomplete Gamma function

$$\Gamma(n, y) = \int_y^\infty dx\, x^{n-1} e^{-x}. \tag{B.5}$$

- The exponential integral

$$\mathrm{Ei}\,(x) = \int_x^\infty du \frac{e^{-u}}{u}. \tag{B.6}$$

  1. Asymptotic behavior

$$\mathrm{Ei}\,(x) = -\ln x - \gamma \tag{B.7}$$

- The Beta function

$$B(n, m) = \int_0^1 dx\, x^{n-1}(1-x)^{m-1}. \tag{B.8}$$

  1. Relation to Gamma function

$$B(n, m) = \frac{\Gamma(n)\Gamma(m)}{\Gamma(n+m)} \tag{B.9}$$