UNIVERSITY OF CALIFORNIA

Los Angeles

A Reversible Interacting Particle System on the Homogeneous Tree

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Mathematics

by

Amber Lynn Puha

1998

The dissertation of Amber Lynn Puha is approved.

Bryan Ellickson

Thomas S. Mountford

Roberto H. Schonmann

Thomas M. Liggett, Committee Chair

University of California, Los Angeles

1998

DEDICATION

Dedicated to Wally, Mom, and Dad.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First and foremost, I acknowledge the effort and attention that Professor Thomas M. Liggett has invested my training as a mathematician. As my thesis advisor, he was persistent and demanding, but no more on me than on himself. His passion for mathematics has been contagious. And, he has set an outstanding example of what a mathematician is and should be. For that, I am forever in his debt.

With great appreciation, I recognize the encouragement that I have received from many others. Professor Jennifer T. Chayes, Professor Thomas S. Mountford, and Professor Roberto H. Schonmann have all been instrumental at various stages in my graduate education. Additionally, none this would have ever been conceived were it not for Professor Ruth J. Williams. During my undergraduate years, she instilled me with the confidence to continue my studies. It is because of these people that I have discovered the pleasures doing of mathematics.

VITA

April 30, 1967	Born, Oceanside, California
1993	B.A., Mathematics University of California, San Diego La Jolla, California
1993	John Muir College Scholar of the Year University of California, San Diego La Jolla, California
1993–1997	Eugene V. Cota Robles Fellow Department of Mathematics University of California, Los Angeles
1993–1997	Teaching Assistant Department of Mathematics University of California, Los Angeles
1995	M.A., Mathematics Department of Mathematics University of California, Los Angeles
Summer 1996	Teaching Assistant IAS/PCMI Graduate Summer Program Institute for Advanced Study Princeton, New Jersey
1997	Robert Sorgenfrey Distinguished Teaching Assistant Department of Mathematics University of California, Los Angeles
1997–1998	Dissertation Year Fellow Department of Mathematics University of California, Los Angeles
1997	Chancellor Charles and Sue Young Graduate Student Awardee

Department of Mathematics
University of California, Los Angeles

1998 University of California
President's Postdoctoral Fellow
Department of Mathematics
University of California, San Diego
La Jolla, California

PUBLICATIONS AND PRESENTATIONS

J. T. Chayes, A. L. Puha, and T. D. Sweet(1996). Independent and Dependent Percolation. *IAS/PCMI Summer Lecture Series,* to appear.

Amber L. Puha (August 1997). *A Reversible Nearest Particle System on the Binary Tree.* Paper presented at the meeting of the Institute of Mathematical Statistics, Anaheim, California.

Amber L. Puha (January 1998). *A Reversible Nearest Particle System on the Binary Tree.* Paper presented at the meeting of the American Mathematical Association, Baltimore, Maryland.

Amber L. Puha (1998). A Reversible Nearest Particle System on the Homogeneous Tree. *Journal of Theoretical Probability,* to appear.

ABSTRACT OF THE DISSERTATION

A Reversible Interacting Particle System

on the Homogeneous Tree

by

Amber Lynn Puha

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 1998

Professor Thomas M. Liggett, Chair

An interacting particle system is a stochastic processes in which particles live on the vertices of some infinite graph. They are created and destroyed over time via local probabilistic rules. One feature of the rules for the system studied here is that if all of the particles are destroyed, then no more particles can ever be created. This leads one to consider the notion of survival whereby the set of particles avoids absorption into the empty set in finite time. More specifically, one wants to know if survival occurs with positive probability.

Another feature of the rules for this system is that the creation of particles is controlled by a growth parameter β. In fact, it turns out to be the case that

the survival probability is an increasing function of β. This has the consequence that there is critical value β_c of the growth parameter. The process survives with positive probability when β is larger than β_c, while the does not survive when β is smaller than β_c . The central objective of this work is to devise a method for obtaining good bounds on β_c. Furthermore, one would like to ascertain how well the process survives when it does so.

Interest in the local rules chosen for study here arises from the fact that they lead to a reversible interacting particle system. Reversibility admits tools applicable to studying the survival properties of the system. For example, the Dirichlet principle can be used to express the survival probability as an infimum of a certain variational functional over all functions in some class. Using these tools, upper and lower bounds are obtained on the critical value β_c of the growth parameter. The results are the sharpest on the binary tree where the bounds are sufficiently good to completely determine the critical value. It is also shown that if the process survives, then it survives in a fairly strong sense. Moreover, these tools are used to obtain estimates on the rate at which the survival probability tends to zero as the growth parameter approaches the critical value β_c.

CHAPTER 1

Introduction

Interacting particle systems are probabilistic models that are defined in terms of local interactions. Such models are useful for modeling physical and biological systems. Often times these systems undergo a sudden change in the long run behavior as some parameter varies. For instance, water held at a temperature below freezing turns to ice after a long period of time, while water held at a temperature slightly above freezing remains liquid indefinitely. The freezing point is what is known as a *critical value* of the temperature parameter. In a biological system, whether or not a given species survives may depend on the reproduction rate, so that the notion of the freezing point can be replaced by the notion of the survival threshold. These abrupt changes in the long run behavior of the system are known as *phase transitions*. Understanding the nature of phase transitions motivates much of the interacting particle systems research. A widely accept notion called universality declares that the behavior of these systems near the critical value is very robust: the near critical behavior should depend very little on the details of the local interaction and should be determined by the universality class to which the system belongs. This suggests that simple mathematical models can provide accurate predictions about the near critical behavior of real systems.

One of the most widely studied interacting particle systems is the *contact process*. Harris [14] introduced the contact process in 1974 to model the spread of an infection through a population. The population structure is represented by a graph with the vertices denoting individuals. Infected individuals (particles) infect their neighbors at rate λ. Once infection sets in, recovery occurs at rate one. On a global scale, the infection either dies out with probability one or persists forever with positive probability. Furthermore, the probability $s(\lambda)$ that the infection persists forever starting with a single infected individual is an increasing function of λ. Therefore, it is natural to define the critical value $\lambda_c = \inf\{\lambda : s(\lambda) > 0\}$. It is not difficult to show that $0 < \lambda_c < \infty$, that is that the contact process undergoes a phase transition. An active area of research concerns findings good bounds on λ_c for various graphs.

The notion of survival makes sense for the *finite system*, the system in which the number of infected individual at any given time is finite. One might ask about the long run behavior of the system when an infinite number of individuals are infected initially. This question leads to the study of invariant measures for the interacting particle system and their domains of attraction. In this context, the interacting particle system undergoes a phase transition if there exist values of the parameter λ for which the the sets of extremal invariant measures have distinct structures. In the case of the contact process on the d-dimensional integer lattice $\mathbb{Z}^d$, $\lambda \leq \lambda_c$ implies that the pointmass on the state with all healthy individuals is the only invariant measure. However, for $\lambda > \lambda_c$, there exists exactly one nontrivial extremal invariant measure. In particular, the critical values for the infinite and finite system coincide. See [17] for more background on these these topics.

Spitzer [32] proposed a natural generalization of the contact process on the one dimensional integer lattice $\mathbb{Z}$ by allowing the infection rates to depend on the distances to the nearest infected individuals. Such systems are known as *nearest particle systems*. Without making any additional assumptions on the rates, little can be said about the behavior of these systems. However, by restricting attention to the nearest particle systems that are reversible, tools become available that lend themselves to the study of phase transitions. Furthermore, understanding the behavior of the reversible class may provide insight into the nonreversible situation. Spitzer's definition of nearest particle systems depends on the fact that the underlying graph is $\mathbb{Z}$. This leads us to ask what a reversible nearest particle system is in higher dimensions, or on other graphs. While there have been some attempts to define and study reversible systems on graphs besides $\mathbb{Z}$ (see [5], [6], and [19]), the theory is not well developed.

Here a reversible interacting particle system called the uniform model is introduced and studied. The process evolves on the homogeneous tree $\mathbb{T}^d$, a graph with no cycles in which each vertex has degree $d + 1$. The infection rate is the same as for the contact process, while the recovery rate is modified: infected individuals are prevented from recovering when at least two neighbors are infected; otherwise, infected individuals recover at rate one. It turns out that this modification leads to a reversible interacting particle system, as we will see in Section 1.4.

Before presenting the treatment of the uniform model, we pause to make some of the aforementioned notions more precise and to give more background on the related processes. Section 1.1 contains the formal definition of an interacting particle system and defines much of the notation that will be used throughout. Section 1.2 summarizes some of what is known about the contact process, and in so doing, introduces the key ideas that motivate the questions that are addressed for the uniform model. Section 1.3 gives the formal definition of a reversible interacting particle system and outlines much of what is known about Spitzer's reversible nearest particle systems. In Section 1.4, we return to the uniform model. Here, the problems that will interest us are described. This section contains the statements of the main theorems that will be proved in the ensuing chapters and a pointer to where each theorem is proved.

1.1 Notation and Preliminaries

Interacting particle systems are continuous time Markov processes. Each system has a spin space S and an underlying graph $G = (V, E)$, where V denotes the set of vertices of the graph and E denotes the set of edges. Typically, vertices are referred to as *sites*. The process takes values in the space $X = S^V$ of all possible labelings of vertices (or sites) of the graph with elements in S. Typically, elements of S are called *spins*, or spin values. In all of the examples discussed here $S = \{0, 1\}$ and the spin value 1 is associated with being infected, or being occupied. An element $\eta \in X$ is referred to as a *configuration*. Configurations can be viewed as functions $\eta : V \to S$ with $\eta(x)$ denoting the spin at site x in configuration η. They can also be viewed as subsets of V via the identification $\eta \leftrightarrow \{y \in V : \eta(y) = 1\}$. For future reference, note that there is a natural partial ordering on X that is given by $\eta \leq \zeta$ if and only if $\eta(x) \leq \zeta(x)$ for all $x \in V$.

The evolution is determined by a nonnegative rate function $c(x, \eta)$ that specifies the rate at which the spin at site x changes (or flips) from $\eta(x)$ to $1 - \eta(x)$ in configuration η. Notice that only one spin value changes in a single transition. Systems with this property are called *spin systems*. It is certainly possible to have models in which more than one spin flips in a single transition. The discussion here is restricted to spin systems and the rate function is often referred to as the collection of *flip rates*.

Given a collection of flip rates, one can define a generator G on a certain dense subset $D(X)$ of the continuous functions $C(X)$ on X:

$$Gf(\eta) = \sum_y c(y, \eta) \Big(f(\eta_y) - f(\eta) \Big) \quad \text{for all } f \in D(X).$$

Here the configuration η_y agrees with η except at y, where it takes the value $1 - \eta(y)$. Under appropriate assumptions on the rates, the generator G uniquely determines a Feller Markov process with state space X. A sufficient condition is that the rates are uniformly bounded and that for all x, $c(x, \cdot)$ depends on the values of η within some finite distance R of x. See [17] for a more detailed account of the construction.

We denote the semigroup and state of the process at time t by $S(t)$ and η_t respectively. Also, P^η is the probability measure that puts mass one on sample paths with $\eta_0 = \eta$. Accordingly, η_t^η is the value of the process at time t when $\eta_0 = \eta$ almost surely. It is instructive to note that the rates are related to the measure P^η by the equation

$$P^\eta\left(\eta_t(x) \neq \eta(x)\right) = c(x, \eta)t + o(t) \qquad \text{as } t \searrow 0.$$

The expected value with respect to the measure P^η is denoted by $\mathbb{E}^\eta$. For μ in the set of all probability measures $\mathcal{P}$ on X, $\mu S(t)$ denotes the distribution of the process at time t when the initial distribution is μ.

1.2 The Contact Process

For the contact process, the rate function is given by

$$c(x, \eta) = \begin{cases} \lambda \sum_{\{y:||x-y||=1\}} \eta(y) & \text{if } \eta(x) = 0, \\ 1 & \text{otherwise.} \end{cases}$$

Here $\lambda > 0$ is the *growth parameter* and $||x - y||$ denotes the length of the shortest path from x to y in the graph. The rates for the contact process satisfy a very useful property known as attractiveness. In order to define an attractive process, say that $f \in C(X)$ is increasing if $f(\eta) \le f(\zeta)$ for all $\eta \le \zeta$. An attractive (or monotone) process obeys the following condition: if f increasing, then $S(t)f$ is increasing for all $t \ge 0$. For spin systems, attractiveness is equivalent to the rate function satisfying

$$c(x,\eta) \le c(x,\zeta) \quad \text{if } \zeta(x) = 0 \qquad \text{and} \qquad c(x,\eta) \ge c(x,\zeta) \quad \text{if } \eta(x) = 1, \tag{1.2.1}$$

for all $\eta \le \zeta$. Typically, this condition is easy to verify and in particular it holds for the contact process. One powerful consequence of attractiveness is that for any $\eta \le \zeta$ there exists a coupling (η_t, ζ_t) of two copies of the process such that

$$P^{(\eta,\zeta)}(\eta_t \le \zeta_t) = 1 \tag{1.2.2}$$

for all $t \ge 0$. In Section 2.1, this coupling will be reviewed.

As a consequence of attractiveness, the process converges in distribution when the initial measure is the pointmass δ_V on the state with all sites occupied. The limiting measure is invariant and stochastically dominates all other invariant measures for the process. To see this, note that $\delta_V \ge \delta_V S(t)$ so that by attractiveness $\delta_V S(s) \ge \delta_V S(t+s)$. In particular, the measures $\delta_V S(t)$ are decreasing in t and therefore converge. The limiting measure ν is called the *upper invariant measure.* The term invariant means that $\nu S(t) = \nu$ for all times t. The measure ν is invariant because it is obtained as a limit: $\nu = \lim_{t\to\infty} \delta_V S(t) = \lim_{t\to\infty} \delta_V S(t+s) = \nu S(s)$. It dominates all other invariant measures because $\delta_V \ge \mu$ implies that $\delta_V S(t) \ge$

$\mu S(t)$. If μ is invariant, then this inequality becomes $\delta_V S(t) \geq \mu$. Letting t tend to infinity gives $\nu \geq \mu$.

The upper invariant measure is intimately connected to the behavior of the finite system. It turns out that ν is nontrivial if and only if the probability of survival is nonzero. Moreover, the finite system often times converges in distribution to a convex combination of $\delta_\emptyset$ and the upper invariant measure. This is a phenomenon known as *complete convergence* and the constant is given by the probability of survival:

$$P^A S(t) \to P^A(\eta_t \neq \emptyset \ \forall \ t)\nu + P^A(\eta_t = \emptyset \text{ some } t)\delta_\emptyset.$$

It is immediate that complete convergence holds whenever $s(\lambda) = 0$. In particular, it holds for $\lambda < \lambda_c$. One of the most important results about the contact process on $\mathbb{Z}^d$ is that complete convergence holds for all λ. Durrett [9] proved a one dimensional version of this result that applies for $\lambda > \lambda_c$ in 1980. A proof can be found in [17]. It relies on the notion of edge speeds which are only defined in one dimension. Next came many partial results that hold in all dimensions, but only for λ sufficiently large; see [10], [30], and [1]. In 1991, Bezuidenhout and Grimmett [4] proved that the contact process obeys complete convergence for all dimensions and all λ. Their work includes a proof that the critical contact process dies out: $s(\lambda_c) = 0$.

To this point, the discussion has centered around a notion of *global survival*, the event that the process is not absorbed into the empty set in finite time. There is also a notion of *local survival*, the event that a given site is infected at an unbounded collection of times. On $\mathbb{Z}^d$, global survival implies local survival as a consequence of complete convergence; however, this is not true for all graphs. Pemantle [25] first observed this fact by investigating the behavior of the contact process on the homogeneous tree. He proved that on these graphs the contact process experiences an intermediate phase, provided that the degree is sufficiently large. This intermediate phase is characterized by nonlocal, global survival. Fix a distinguished vertex O in the tree that will be referred to as the origin, or the root. Let

$$\lambda_2 = \inf\{\lambda : P^O\,(\eta_t \neq \emptyset \ \forall\ t\ ) > 0\} \text{ and}$$

$$\lambda_3 = \inf\{\lambda : P^O\,(O \in \eta_t \text{ for unbounded } t\ ) > 0\}.$$

Notice that $\lambda_2 = \lambda_c$. Global survival without local survival is called *weak survival*. In particular, weak survival occurs with positive probability if $\lambda \in (\lambda_2, \lambda_3)$. By obtaining upper bounds on λ_2 and lower bounds on λ_3, Pemantle showed that $\lambda_2 < \lambda_3$ for $d \geq 3$, leaving the case $d = 2$ open. This case was handled by Liggett [21] using a similar but more sophisticated approach. Shortly after, Stacey [33] came up with a proof that $\lambda_2 < \lambda_3$ that did not yield explicit bounds, but worked for all $d \geq 2$.

A central feature of these arguments involves finding a suitable function of the state space that is a nonnegative supermartingale with respect to the evolution of the contact process. Two supermartingales turned out to be extremely useful. In order to define these two supermartingales, a level $\ell(x)$ is assigned to each vertex x in the tree such that $\ell(O) = 0$. The level is assigned inductively to the $d+1$ neighbors of the vertex x by setting the level of one of the neighbors equal to $\ell(x) - 1$ and the level of the other d neighbors equal to $\ell(x) + 1$. Let $0 < \rho < 1$ and define

$$f_\rho(A) = \rho^{|A|} \quad \text{and} \quad g_\rho(A) = \sum_{x \in A} \rho^{\ell(x)}, \tag{1.2.3}$$

for finite $A \in X$. Notice that if $f_\rho(\eta_t)$ converges to zero, then $|\eta_t|$ tends to infinity. Furthermore, if $g_\rho(\eta_t)$ tends to zero, then the process dies out locally. These observations indicate why these functions would be good candidates. The game is to determine for which λ these functions are in fact supermartingales for some ρ. As the degree decreases, it becomes more difficult to obtain bounds sufficiently good to separate λ_2 and λ_3. Pemantle's original theorem was that

$$\lambda_2 \leq \frac{1}{d-1} \quad \text{and} \quad \frac{1}{2\sqrt{d}} \leq \lambda_3,$$

which is enough to prove that $\lambda_2 < \lambda_3$ provided $d \geq 6$. He then went on to improve the bounds enough to separate λ_2 and λ_3 for $d > 2$. Eventually, Liggett showed that $\lambda_2 \leq .605$ and $.609 \leq \lambda_3$ for $d = 2$ which barely separates the two. Soon after, Stacey discovered an alternative approach involving the function g_ρ that did not produced bounds, but that did show the existence of an intermediate phase for all $d \geq 2$.

Later it was observed by Liggett [20] that $\left(\mathbb{E}^O(g_\rho(\eta_t))\right)^{1/t}$ converges to a finite limit. He initiated a study of this limiting function $\phi(\lambda, \rho)$ which depends on both

λ and ρ. One objective of this work was to describe the behavior of the function

$$u(n) = P^O(x_n \in \eta_t \text{ some } t\ ) \qquad \text{as } n \to \infty,$$

where $\|x_n - O\| = n$. This probability turns out to decay exponentially if $\phi(\lambda, \rho) < 1$ for some ρ (which holds if and only if $\lambda < \lambda_3$). It is even possible to determine the critical rates of decay:

$$u(n)^{1/n} \nearrow \frac{1}{d} \qquad \text{at } \lambda_2 \qquad \text{and} \qquad u(n)^{1/n} \nearrow \frac{1}{\sqrt{d}} \qquad \text{at } \lambda_3.$$

It is immediate that there is no decay if $\lambda > \lambda_3$. The critical rate of decay at λ_2 comes from combining an upper bound in [20] and results of Schonmann [31]. The critical rate of decay at λ_3 was originally conjectured to be an upper bound by Liggett. Lalley and Sellke [15] proved Liggett's conjecture and ultimately, their proof was simplified by Schonmann and Salzano [28]. A good place to read about this work is [22].

An important property of the contact process that is essential to the aforementioned analysis of $g_\rho(\eta_t)$ is that it is additive. This means that there exists a coupling such that

$$\eta_t^A = \cup_{x \in A} \eta_t^x \quad \text{for all } t \geq 0. \tag{1.2.4}$$

The existence of this coupling comes from a graphical construction of the contact process that we will not go over here. Instead the reader is referred to [22]. Suffice it to say that additivity implies that $\mathbb{E}g(\eta_t^A) = \mathbb{E}g(\cup_{x \in A} \eta_t^x) \leq \sum_{x \in A} \mathbb{E}g(\eta_t^x)$, which is both a first step toward constructing a supermartingale and a first step towards showing that $\left(\mathbb{E}^O(g_\rho(\eta_t))\right)^{1/t}$ converges to a finite limit.

In the intermediate phase complete convergence must fail. Here the infection is 'wandering off to infinity' so to speak and therefore the finite process is converging in distribution to $\delta_\emptyset$. This peaked interest in the set of invariant measures for the contact process on $\mathbb{T}^d$ in the intermediate phase. Durrett and Schinazi [11] discovered that there are infinitely many extremal invariant measures in this phase. These measures have the property that the density of particles tends to a nonzero constant on some significant proportion of the boundary of the tree. Liggett [20] also produced a spherically symmetric collection of invariant measures for which the density of particles tends to zero near the boundary.

By this time, Zhang [35] had proved that the contact process on $\mathbb{T}^d$ obeys complete convergence for $\lambda > \lambda_3$. It is not obvious from the definition of complete convergence that survival together with complete convergence is a monotone increasing property of λ. Nevertheless, it turned out to be the case on both $\mathbb{Z}^d$ and $\mathbb{T}^d$. This interested Schonmann and Salzano ([27] and [29]) who undertook a study of the contact process on arbitrary graphs with monotonicity of complete convergence in mind. One consequence of their work was a criterion for survival together with complete convergence on homogeneous graphs that is obviously monotone increasing in λ:

$$\lim_{n\to\infty} \liminf_{t\to\infty} P^{B(O,n)}(\eta_t \cap B(O,n) \neq \emptyset) = 1,$$

where $B(O,n)$ is the ball centered at the origin of radius n. This leads to the definition of yet another critical value

$$\lambda_4 = \inf\{\lambda : \lim_{n\to\infty} \liminf_{t\to\infty} P^{B(O,n)}(\eta_t \cap B(O,n) \neq \emptyset) = 1\}. \tag{1.2.5}$$

In these terms, the Zhang result states that $\lambda_3 = \lambda_4$. For a proof that $\lambda_3 = \lambda_4$, the reader is referred to [28].

Turning our attention to the near critical behavior of the contact process on $\mathbb{T}^d$, if

$$\lim_{\lambda \searrow \lambda_2} \frac{\log P^O(\eta_t \neq \emptyset \ \forall\ t)}{\log(\lambda - \lambda_2)^{\beta}} = 1,$$

then the survival probability decays like a power law with exponent β. This exponent β is said to be the *critical exponent* of the survival probability. For the contact process on the homogeneous tree, Barksy and Wu [3] showed that if a condition called the triangle condition holds, then the exponent takes its mean field value which is one. Wu [34] verified that this condition holds for $d \geq 5$. Later Schonmann [31] completed the story by verifying that this condition holds for $d \geq 2$. Another quantity that typically displays power law behavior is the expected total space time occupation measure, or the *susceptibility*. It is defined as

$$\chi(\lambda) = \mathbb{E}^O \int_0^\infty |\eta_t| \mathrm{d}t.$$

Letting τ be the time of absorption into the empty set, we see that

$$\mathbb{E}^O \tau \leq \chi(\lambda),$$

and consequently $\mathbb{E}^O \tau = \infty$ implies that $\chi(\lambda) = \infty$. The question then becomes at what rate does $\chi(\lambda)$ diverge. Let

$$\lambda_1 = \inf\{\lambda : \mathbb{E}^O \tau = \infty\}.$$

For the contact process on $\mathbb{T}^d$, it is known that $\lambda_1 = \lambda_2$ and that

$$\lim_{\lambda \nearrow \lambda_2} \frac{\log \chi(\lambda)}{\log(\lambda - \lambda_2)^{-\gamma}} = 1,$$

for $\gamma = 1$. Here again $\gamma = 1$ is the mean field value. The fact that the exponent exists and takes mean field value follows from the fact that the triangle condition holds as verified by Wu [34] for $d \geq 5$ and Schonmann [31] for $d \geq 2$.

1.3 Reversible Nearest Particle Systems

A finite interacting particle system is said to be *reversible* if there exists a measure π supported on the states with finitely many infected individuals such that

$$\pi(A)c(x, A) = \pi(A \cup x)c(x, A \cup x) \tag{1.3.1}$$

for all states A (except possibly a single absorbing state) with finitely many infected sites and all $x \notin A$. The equations in (1.3.1) are known as the detailed balance equations. If there is no exceptional state, then the detailed balance equations are equivalent to self-adjointness of the operator $S(t)$ with respect to the measure π.

Given the definition of a reversible interacting particle system, the next objective is to give a more formal description of a reversible nearest particle system. As previously mentioned, a nearest particle system on $\mathbb{Z}$ has a rate function of the form

$$c(x,\eta) = \begin{cases} f(l_x(\eta), r_x(\eta)) & \text{if } \eta(x) = 0, \\ 1 & \text{otherwise.} \end{cases}$$

Here $f : \mathbb{N}_+ \times \mathbb{N}_+ \to \mathbb{R}_+$ with $\mathbb{N}_+ = \{1, 2, 3, \dots\}$ and $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$. Also, $l_x(\eta)$ (resp. $r_x(\eta)$) is the distance to the nearest particle to the left (resp. right) of x in configuration η. The function f satisfies some mild conditions that prevent explosions and has the properties that $f(l,\infty) = f(\infty, l) > 0$ and $f(\infty,\infty) = 0$. In particular, the empty set is an absorbing state. Requiring the detailed balance equations to hold (except at the empty set) for some measure π is equivalent to the function f taking the form

$$f(l,r) = \frac{\beta(l)\beta(r)}{\beta(l+r)}$$

for some strictly positive function $\beta(\cdot)$ on $\mathbb{N}_+$. This function $\beta(\cdot)$ is also assumed to satisfy $\sum_{n=1}^{\infty} \beta(n) < \infty$. In case $r = \infty$, $f(l,\infty) = \beta(l)$. The measure π is given (up to constant multiples) by $\pi(x) = 1$ for all $x \in \mathbb{Z}$ and

$$\pi(A) = \prod_{i=0}^{n-1} \beta(x_{i+1} - x_i)$$

for $A = \{x_0, \dots, x_n\}$ with $x_i < x_{i+1}$ for $0 \leq i \leq n-1$ and $n \geq 1$.

In contrast to the contact process, critical values for both the finite and infinite reversible nearest particle systems can be computed exactly. This is achieved by taking advantage of the additional tools that become available in the reversible setting (e.g. the Dirichlet principle for the finite system) that will be discussed at some length in Sections 2.7 and 4.1. Let $\lambda = \sum_{n=1}^{\infty} \beta(n)$.

Theorem 1.3.1 *The finite process survives if and only if* $\lambda > 1$. *Furthermore,*

$$\frac{\lambda - 1}{\lambda} \leq P^0(\eta_t \neq \emptyset \ \forall \ t) \leq \left| \lambda \log \left(\frac{\lambda - 1}{\lambda} \right) \right|^{-1} \quad \textit{for } \lambda > 1.$$

Thus, not only has the critical value been computed exactly; the theorem provides estimates on the rate at which the survival probability decays as λ decreases to λ_2.

Originally, Spitzer [32] was interested in the infinite system, the system in which $\sum_{x \leq 0} \eta(x) = \sum_{x \geq 0} \eta(x) = \infty$. In this context, a measure is said to be reversible if the semigroup is self-adjoint with respect to that measure for all times. Reversible measures are also invariant measures. In order to state the simplest version of the results about reversible measures, the function $\beta(\cdot)$ is assumed to satisfy

$$\frac{\beta(n)}{\beta(n+1)} \searrow 1. \tag{1.3.2}$$

It turns out that the monotonicity in assumption (1.3.2) is equivalent to saying that the system is attractive. For the infinite system, the fact that the limit in (1.3.2) is taken to be one is not a restriction since the function $\beta(\cdot)$ can be replaced with $\beta(\cdot)\alpha^{\cdot}$ without affecting the flip rates. Suppose that there exists a θ such that

$$\sum_{l=1}^{\infty} \beta(l)\theta^l = 1 \qquad \text{and} \qquad \sum_{l=1}^{\infty} l\beta(l)\theta^l < \infty. \tag{1.3.3}$$

In particular, this is the case whenever $\lambda > 1$ and this not the case whenever $\lambda < 1$. If assumption (1.3.3) holds, then $\beta(\cdot)\theta^{\cdot}$ is a strictly positive probability distribution on $\mathbb{N}_+$ that determines a stationary renewal measure μ_β on $\mathbb{Z}$ with increments that are distributed according to this measure.

Theorem 1.3.2 *Assume that (1.3.2) holds. Then the upper invariant measure is $\delta_\emptyset$ whenever there exists no θ satisfying assumption (1.3.3). If assumption (1.3.3) is satisfied for some θ, then the stationary renewal measure μ_β is both the upper invariant measure and the unique nontrivial reversible measure for this process. In particular, the critical values for the finite and infinite system coincide.*

In view of Theorem 1.3.2, one might like to know if there are other invariant measures that are not reversible. Under the additional assumption that

$$\sum_n \frac{\beta(n)\beta(n)}{\beta(2n)} < \infty, \tag{1.3.4}$$

Liggett [16] proved that μ_β is the unique nontrivial translation invariant, invariant measure. This condition turns out to hold for most $\beta(\cdot)$ of interest. Mountford [24] later built on this result and proved a complete convergence theorem under assumption (1.3.4). His proof takes advantage of ideas introduced in [4].

The subcritical approach to critical for reversible nearest particle systems is very well understood. In fact, as a consequence of reversibility, the expected extinction time and susceptibility are exactly computable as functions of λ. One consequence of these computations is that $\lambda_1 = \lambda_2$.

Theorem 1.3.3 *Assume that $\lambda < 1$. Then*

$$\mathbb{E}^0(\tau) = (1-\lambda)^{-1} \qquad \textit{and} \qquad \chi(\lambda) = (1-\lambda)^{-2}.$$

The proofs of all of the theorems stated in this section can be found in [17], with the exception of the complete convergence result. For that the reader is referred to the original paper [24]. Theorems 1.3.1 and 1.3.3 are due to Griffeath and Liggett [13]. Theorem 1.3.2 follows from work of both Spitzer [32] and Liggett [16].

Two important obstacles prevent a direct generalization of reversible nearest particles to graphs other than $\mathbb{Z}$. Firstly, on what quantity should the rate at which a vacant site becomes occupied depend; that is, how should one generalize the notion of the nearest particle to the left and right? Secondly, there is no generalization of a renewal measure even to $\mathbb{Z}^d$ for $d \geq 2$. Liggett [18] introduced what he called the uniform model in an effort to extend the theory of reversible nearest particle systems to $\mathbb{T}^d$. It has two parameters rather than one and henceforth will be referred to as the two parameter uniform model. Some other work in this direction includes [5], [6], and [19].

1.4 The Uniform Model

The flip rates for the uniform model are given by

$$c(x,\eta) = \begin{cases} \beta \sum_{\{y:\|x-y\|=1\}} \eta(y) & \text{if } \eta(x) = 0, \\ 1 & \text{if } \eta(x) = 1 \text{ and } \sum_{\{y:\|x-y\|=1\}} \eta(y) \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

As for the contact process, $\beta > 0$ is the growth parameter and $\|x - y\|$ denotes the length of the shortest path connecting x and y. These dynamics can be viewed as a modification of the contact process: there the rate at which an occupied site becomes vacant is one regardless of the spin values in the neighborhood. Notice that these rates also satisfy (1.2.1). In particular, the uniform model is attractive.

One effect of the modification is that connected components remain connected until absorption into the empty set. To see this, note that the rate at which a 0 flips to a 1 is zero unless there are some 1's in the neighborhood of the 0. Thus, new connected components cannot appear due to a birth. Also note that, if a 1 flips to 0 at a positive rate, then at most one of neighbors of the 1 is occupied. Therefore, anything in the connected component of this 1 is connected through this occupied neighbor. In particular, removing this 1 cannot break the connected component into two pieces. This leads to the observation that the configuration with all sites occupied is absorbing so that the measure $\delta_{\mathbb{T}^d}$ is the upper invariant for the process.

Another important distinction between the two processes is that the uniform model is reversible. Consider the measure $\pi(A) = \beta^{|A|}$ for all finite, connected $A \subset \mathbb{T}^d$, where $|A|$ denotes the number of vertices in the set A. If A is connected and nonempty, then $c(x, A) = \beta$ if and only if $c(x, A \cup x) = 1$. Therefore,

$$c(x, A)\pi(A) = c(x, A \cup x)\pi(A \cup x) \qquad \text{for all } x \notin A.$$

It is easy to see that the contact process is not reversible since isolated particles die at positive rate and cannot be reborn at positive rate until some neighboring vertex becomes occupied.

A third fundamental difference between these two processes is that the uniform model is not additive. In fact, it is superadditive in the sense that there exists a coupling such that

$$\cup_{x \in A} \eta_t^x \subseteq \eta_t^A,$$

and no such coupling holds with an equality. This has the consequence that the supermartingales that were so useful for analyzing the behavior of the contact process on the tree are not supermartingales for the uniform model. Therefore, the techniques used to prove results about this process will differ greatly from those used for the contact process. The methods take advantage of reversibility and tend to be more like those used to prove results about reversible nearest particle systems.

Liggett [18] first introduced the two parameter version of this process in 1985. It has both an interior growth parameter λ and an exterior growth parameter

$\gamma < 1/d$. Given a configuration η, let $\mathcal{G}(\eta)$ be the minimal connected subgraph of $\mathbb{T}^d$ containing η. The rate at which a vacant site becomes occupied in configuration η decays exponentially with the distance to $\mathcal{G}(\eta)$, while occupied sites become vacant at rate one. The flip rates are given by

$$c(x, \eta) = \lambda \gamma^{\|x - \mathcal{G}(\eta)\|}(1 - \eta(x)) + \eta(x),$$

where $\|x - \mathcal{G}(\eta)\| = \min\{\|x - y\| : y \in \mathcal{G}(\eta)\}$. The two parameter model is reversible with respect to the measure $\mu(A) = \gamma^{|\mathcal{G}(A)|}\lambda^{|A|}$ for finite $A \subset \mathbb{T}^d$. Liggett studied the survival properties of the finite system and gave bounds on the critical value of the interior growth parameter in terms of the exterior growth parameter. The connection between the single and double parameter models is that the single parameter uniform model can be regarded as a limit of the double parameter version. To see this, set the double parameter nearest neighbor birth rate $\lambda\gamma$ constantly equal to β while letting the exterior growth parameter γ tend to zero. In particular, the rate at which vacant sites at a distance strictly greater than one from $\mathcal{G}(\eta)$ become occupied tends to zero. Since the interior growth parameter $\lambda = \beta/\gamma$, the interior birth rate tends to infinity. Thus, any occupied site in the interior of $\mathcal{G}(\eta)$ that becomes vacant is instantaneously reoccupied.

The uniform model is also closely related to another process that arises in the computer science literature. In this arena, binary search trees are a common way to store and retrieve data. The search process is modeled by a Markov chain that evolves by adding a vertex at random to the current state. This vertex is chosen from those at distance one from the current state. The state space for the search process is not quite subsets of $\mathbb{T}^2$. It is actually subsets of a subgraph $\mathbb{B}^2$ of $\mathbb{T}^2$. This subgraph $\mathbb{B}^2$ contains a distinguished vertex called the root that has degree two. All other vertices have degree three. This distinguished vertex is the first vertex to be occupied with probability one. Such a process can be view as a discrete time uniform model without deaths. The main interest is the asymptotic height of these trees which was studied independently by Pittel [26] and Devroye [7]. Let H_n be the distance to the farthest vertex from the origin that is in the current state after n additions. Also, let h_n be the distance to the nearest vertex to the origin that is not in the current state after n additions. Both of these quantities grow like $\ln n$, but with different constant rates: With probability one

$$\lim_{n\to\infty} \frac{h_n}{\ln n} = L_1 \qquad \text{and} \qquad \lim_{n\to\infty} \frac{H_n}{\ln n} = L_2,$$

where L_1 and L_2 are the two distinct roots of the equation $L\exp((1-L)/L) = 2$. Barlow, Pemantle, and Perkins [2] later studied more general versions of pure growth processes on trees.

A major objective of this work is to exploit reversibility to provide a complete analysis like that available for reversible nearest particle systems. Motivated by the contact process on $\mathbb{T}^d$, we consider the following critical values of the birth parameter. As before, τ denotes the time of absorption into the empty set and O

is a distinguished vertex referred to as the origin, or the root. Let

$$\beta_1(d) = \inf\{\beta : \mathbb{E}^O\,(\tau) = \infty\},$$

$$\beta_2(d) = \inf\{\beta : P^O\,(\eta_t \neq \emptyset \;\forall\; t\;) > 0\},$$

$$\beta_3(d) = \inf\{\beta : P^O\,(O \in \eta_t \text{ for unbounded } t\;) > 0\}, \text{ and}$$

$$\beta_4(d) = \inf\{\beta : \lim_{n\to\infty} \liminf_{t\to\infty} P(O \in \eta_t^{B(O,n)}) = 1\}.$$

It is immediate that $\beta_1(d) \leq \beta_2(d) \leq \beta_3(d) \leq \beta_4(d)$. The definition of $\beta_4(d)$ is a modified version of definition (1.2.5) that defined λ_4 for the contact process. In Section 2.2, will we see that $\beta_4(d)$ is the threshold for survival and complete convergence. For the contact process on $\mathbb{T}^d$, we know that $\lambda_1 = \lambda_2 < \lambda_3 = \lambda_4$, provided $d \geq 2$. The process is said to be *subcritical* if $\mathbb{E}^O\,(\tau) < \infty$ and *supercritical* if $\lim_{n\to\infty} \liminf_{t\to\infty} P(O \in \eta_t^{B(O,n)}) = 1$.

Theorem 1.4.1 summarizes the main results regarding critical values for the (single parameter) uniform model. On the binary tree, all critical values are computed exactly paralleling the results for reversible nearest particle systems on $\mathbb{Z}$. The analysis itself is similar in the sense that it uses certain tools associated with reversibility. However, many new ideas are required in order to obtain these results. The proof of this theorem is the subject of Chapter 2.

Theorem 1.4.1

a) For $d \geq 2$, $\beta_2(d) = \beta_3(d)$.

b) For $d \geq 2$,

$$\beta_1(d) = \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}.$$

Furthermore, at $\beta_1(d)$ *the expected extinction time is finite.*

c) For $d \geq 2$,

$$\beta_4(d) \leq \frac{d}{2(d-1)^2}.$$

d) For $d = 2$, $\beta_1(2) = \beta_4(2) = \frac{1}{4}$.

Theorem 1.4.1a) states that in contrast to the contact process on $\mathbb{T}^d$ the uniform model has no intermediate phase characterized by weak survival for all $d \geq 2$. The key factor that will be taken advantage of in the proof of Theorem 1.4.1a) is connectedness of the uniform model. By b), $\beta_1(d)$ is asymptotically $1/ed$ and the bound given in c) is asymptotically $1/2d$. These values are close, but not close enough to rule out an intermediate phase. Part d) states that there is in fact no intermediate phase in $d = 2$ and identifies the exact location of the phase transition.

The technique used to push the upper bound on $\beta_4(2)$ down to $\beta_1(2)$ may work for general d. The remaining obstacle is to show that a certain set of equations has a solution that is absolutely bounded by one (see Lemma 2.9.1). A limiting version of these equations yields a partial differential equation. This PDE does in fact have a solution that is absolutely bounded by one.

Theorem 1.4.2 *For $d \geq 3$, let $\alpha^* : \mathbb{R}_+^d \to \mathbb{R}$ be defined by*

$$\alpha^*(x_1, \ldots, x_d) = \sum_{i=2}^{d} \frac{(x_1 - x_i)(x_1^2 + 10x_1x_i + x_i^2)(x_2 + \cdots + x_{i-1} + x_{i+1} + \cdots + x_d)}{d(d-2)(x_1 + \cdots + x_d)(x_1 + x_i)^3} + \frac{1}{d}.$$

Then $\alpha^*(x_1, \ldots, x_d)$ is symmetric in the variables $x_2, \ldots, x_d$, absolutely bounded by one, and a solution to

$$\sum_{i=1}^{d} \alpha^*(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) = 1$$
$$\sum_{i=1}^{d} \left(\frac{3}{2x_i} - \frac{\partial}{\partial x_i} \right) \alpha^*(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) = \frac{3}{2(x_1 + \cdots + x_d)}.$$

The analysis of the PDE that is presented in Chapter 3 may be of independent interest. Firstly, the PDE relates values of the function and its derivatives at distinct (not necessarily close) points in the positive orthant. Furthermore, simple inspection of the PDE does not suggest a particular form for a candidate solution. Therefore, some strategy must be implemented in order to find the solution exhibited in Theorem 1.4.2.

Theorem 1.4.2 suggests that the next conjecture holds. The conjecture implies that the uniform model undergoes exactly one phase transition on all homogeneous trees.

Conjecture 1.4.3 *For $d \geq 3$, $\beta_4(d) = \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$.*

In Chapter 4, our attention turns to the behavior near the critical value $\beta_1(d)$ which, as the conjecture indicates, we believe to be the only critical value for the uniform model. We prove the following theorem.

Theorem 1.4.4 *On* $\mathbb{T}^d$*,*

$$\limsup_{\beta \searrow \beta_1} \frac{P^O(\eta_t \neq \emptyset \ \forall \ t)}{(\beta - \beta_1)^{5/2}} \leq C_1 \tag{1.4.1}$$

for some constant $0 < C_1 < \infty$*. On the binary tree,*

$$\liminf_{\beta \searrow 1/4} \frac{P^O(\eta_t \neq \emptyset \ \forall \ t)}{(\beta - 1/4)^{1+\sqrt{13}/2}} \geq C_2 \tag{1.4.2}$$

for some constant $0 < C_2 < \infty$*.*

One might be tempted to conclude that inequality (1.4.1) implies that the survival probability is continuous at the critical value $\beta_1(d)$. However, this follows from other considerations. There is relatively elementary argument that shows that the survival probability is right continuous that will be given in Section 4.3. Since Theorem 1.4.1a) states that the survival probability is zero at $\beta_1(d)$, this together with right continuity implies continuity at $\beta_1(d)$. For $d \geq 3$, inequality (1.4.1) is only interesting if $\beta_1(d) = \beta_2(d)$, which we believe to be true. Together the two inequalities in the theorem give bounds on the rate at which the survival probability tends to zero as the growth parameter tends to $\beta_2(2) = 1/4$. The theorem says that on the binary tree the critical exponent for the survival probability of the uniform model lies in the interval $[5/2, 1 + \sqrt{13}/2]$ (if it exists). For $d \geq 3$, a similar result may hold. The main obstacle is proving that the hypothesis of Lemma 2.9.1 are satisfied. This is explained more fully at the end of in Chapter 4.

In Chapter 5, the collection of extremal reversible measures is identified. In order to parameterize this collection of measures, we need to introduce the notion of a backbone in $\mathbb{T}^d$. Say that $b \subseteq \mathbb{T}^d$ is a *backbone* if b is nonempty and for all $x \in b$, there exist at least two neighbors of x that are also in b.

Theorem 1.4.5 *For $\beta \leq \beta_1(d)$, $\{b \subseteq \mathbb{T}^d : b \text{ is a backbone}\} \cup \emptyset$ is in one-to-one correspondence with the collection of extremal reversible measures.*

This correspondence is made precise in Chapter 5. Theorem 1.4.5 suggests many open questions. For example, it would be interesting to determine if there are any invariant measures that are not reversible. Another possibly more difficult question is to determine the domain of attraction for each of these measures. A preliminary step in that direction might be to begin with a product measure at density p and to determine for which p the process converges in distribution to $\delta_{\mathbb{T}^d}$. An easy upper bound on the critical p is the percolation threshold for independent site percolation on $\mathbb{T}^d$, which is $1/d$. It would even be interesting to show that the process converges in distribution to $\delta_\emptyset$ for p in a sufficiently small neighborhood of zero. This would establish that there is a phase transition for fixed $\beta \leq \beta_1(d)$ as p varies.

CHAPTER 2

Critical Values

The critical values $\beta_1(d)$, $\beta_2(d)$, $\beta_3(d)$, and $\beta_4(d)$ signify abrupt changes in strength with which the uniform model survives. We have already observed that $\beta_1(d) \leq \beta_2(d) \leq \beta_3(d) \leq \beta_4(d)$. This raises the question as to whether or not any of these inequalities is strict. One approach to resolving this issue is to devise a method for finding bounds sufficiently good to separate two critical values, or sufficiently good to prove that they are equal. The main portion of this chapter is is devoted to obtaining such estimates and thereby proving Theorem 1.4.1.

Before proceeding to obtain these estimates, we investigate the behavior of the uniform model in the survival regime, $\{\beta : P^O(\eta_t \neq \emptyset \ \forall \ t) > 0\}$. In Section 2.2, qualitative differences in the nature of the survival from one phase to the next are described. These notions apply quite generally. Section 2.3 is devoted to the proof of part a) of Theorem 1.4.1. Here basic properties of the uniform model are exploited to rule out weak survival. The remainder of the chapter is dedicated to proving parts b), c), and d) of Theorem 1.4.1. Reversibility plays an essential role in many of the proofs.

2.1 Some Background on Coupling and Positive Correlations

In the introduction, it was noted that the rates for the uniform model satisfy

$$c(x,\eta) \le c(x,\zeta) \quad \text{if } \zeta(x) = 0 \quad \text{and} \quad c(x,\eta) \ge c(x,\zeta) \quad \text{if } \eta(x) = 1, \tag{2.1.1}$$

for $\eta \le \zeta$. An important consequence of this is that there exists a coupling of two copies of the uniform model so that $\eta \le \zeta$ implies that

$$P^{(\eta,\zeta)}(\eta_t \le \zeta_t) = 1, \tag{2.1.2}$$

for all $t \ge 0$. From this coupling, it readily follows that if f is increasing, then $S(t)f$ is also increasing for all $t \ge 0$. Equivalently, if two probability measures on X satisfy $\mu_1 \le \mu_2$, then $\mu_1 S(t) \le \mu_2 S(t)$ for all $t \ge 0$. A process that satisfies these two equivalent conditions is said to be *attractive.*

The coupled process has the following rates: For $\eta \le \zeta$, let

$$(\eta,\zeta) \to (\eta_x,\zeta_x) \qquad \text{at rate} \qquad \begin{cases} c(x,\eta) & \text{if } \eta(x) = \zeta(x) = 0, \\ c(x,\zeta) & \text{if } \eta(x) = \zeta(x) = 1, \end{cases}$$

$$(\eta,\zeta) \to (\eta,\zeta_x) \qquad \text{at rate} \qquad \begin{cases} c(x,\zeta) - c(x,\eta) & \text{if } \eta(x) = \zeta(x) = 0, \\ c(x,\zeta) & \text{if } \eta(x) = 0, \zeta(x) = 1, \end{cases}$$

$$(\eta,\zeta) \to (\eta_x,\zeta) \qquad \text{at rate} \qquad \begin{cases} c(x,\eta) - c(x,\zeta) & \text{if } \eta(x) = \zeta(x) = 1, \\ c(x,\eta) & \text{if } \eta(x) = 0, \zeta(x) = 1. \end{cases}$$

For $\eta \not\le \zeta$, the two processes evolve independently. The fact that these rates are nonnegative follows from condition (2.1.1). Observe that if $\eta \le \zeta$ and $(\eta,\zeta) \to$

(ϑ, ξ) at positive rate, then $\vartheta \leq \xi$ which guarantees that equation (2.1.2) holds. Also note that $\eta \to \eta_x$ at rate $c(x,\eta)$ and $\zeta \to \zeta_x$ at rate $c(x,\zeta)$ so that the marginal processes both have the same distribution as the original spin system.

One very useful fact about attractive spin systems is that the evolution preserves positive correlations. More precisely, if μ has positive correlations in the sense that

$$\int fh\mathrm{d}\mu \geq \int f\mathrm{d}\mu \int h\mathrm{d}\mu$$

for all increasing functions f and h in $C(X)$, then $\mu S(t)$ also has positive correlations. This holds as a consequence of Harris' Theorem:

Theorem 2.1.1 *Suppose that $S(t)$ and G are respectively the semigroup and generator of an attractive Feller process on X. Assume further that G is a bounded operator. Then the following statements are equivalent:*

$$Gfh \geq fGh + hGf \text{ for all increasing } f, h \in C(X). \tag{2.1.3}$$

$$\mu S(t) \text{ has positive correlations whenever } \mu \text{ does.} \tag{2.1.4}$$

A proof Harris' Theorem can be found in [17]. In order to see that condition (2.1.3) holds for any spin system, note that

$$(f(\eta_x) - f(\eta))\,(h(\eta_x) - h(\eta)) \geq 0,$$

whenever f and h are increasing. This together with the observation that

$$G(fh)(\eta) - f(\eta)Gh(\eta) - h(\eta)Gf(\eta) \qquad (2.1.5)$$

$$= \sum_{x \in V} c(x,\eta)\,(f(\eta_x) - f(\eta))\,(h(\eta_x) - h(\eta))$$

implies that condition (2.1.3) holds.

Harris' Theorem has the following corollary, a proof of which can also be found in [17]:

Corollary 2.1.2 *Suppose that the assumptions of Theorem 2.1.1 are satisfied, and that equivalent conditions (2.1.3) and (2.1.4) also hold. Let η_t be the corresponding process, where the distribution of η_0 has positive correlations. Then for $t_1 < \cdots < t_n$, the joint distribution of $(\eta_{t_1}, \ldots, \eta_{t_n})$, which is a probability measure on X^n, has positive correlations.*

The state space for the coupled process X^2 also has a natural partial ordering: $(\eta, \zeta) \leq (\vartheta, \xi)$ if $\eta \leq \vartheta$ and $\zeta \leq \xi$. Therefore, the notions of attractive processes and positive correlations apply in this context. For the coupled process (η_t, ζ_t) with $\eta_0 \leq \zeta_0$, equation (2.1.2) implies that there are three possible spin values at any given site: $(0,0)$, $(0,1)$, and $(1,1)$. In light of this observation, one can make the identification

$$(0,0) \leftrightarrow 0, \qquad (0,1) \leftrightarrow 1, \qquad \text{and} \qquad (1,1) \leftrightarrow 2,$$

and take the spin space to be $\{0,1,2\}$. Configurations in the space $\{0,1,2\}^V$ will be distinguished from configurations in $\{0,1\}^V$ with an overbar as follows: $\bar{\eta} \in \{0,1,2\}^V$. The aforementioned partial ordering can be expressed as $\bar{\eta} \leq \bar{\zeta}$ if $\bar{\eta}(x) \leq \bar{\zeta}(x)$ for all $x \in V$. Let $c_i(x,\bar{\eta})$ be the rate at which the spin at x flips to i in configuration $\bar{\eta}$. Also, let (η,ζ) be corresponding representation of $\bar{\eta}$ in the space X^2. In particular, let $\eta(x) = 1$ if $\bar{\eta}(x) = 2$; otherwise, $\eta(x) = 0$. Also, let $\zeta(x) = 0$ if $\bar{\eta}(x) = 0$; otherwise, $\zeta(x) = 1$. We have

$$\begin{aligned}
c_0(x,\bar{\eta}) &= c(x,\zeta) && \text{if } \bar{\eta}(x) \neq 0, \\
c_1(x,\bar{\eta}) &= \begin{cases} c(x,\zeta) - c(x,\eta) & \text{if } \bar{\eta}(x) = 0, \\ c(x,\eta) - c(x,\zeta) & \text{if } \bar{\eta}(x) = 2, \end{cases} \\
c_2(x,\bar{\eta}) &= c(x,\eta) && \text{if } \bar{\eta}(x) \neq 2.
\end{aligned} \tag{2.1.6}$$

Here it is understood that $c_i(x,\bar{\eta}) = 0$ in the remaining cases.

The first objective is to show that the coupled process satisfies condition (2.1.3). In this context, (2.1.5) becomes

$$\begin{aligned}
G(fh)(\bar{\eta}) - f(\bar{\eta})Gh(\bar{\eta}) - h(\bar{\eta})Gf(\bar{\eta}) \\
= \sum_{x\in V} \sum_{i=0}^{2} c_i(x,\bar{\eta}) \left(f(\bar{\eta}_x^i) - f(\bar{\eta})\right) \left(h(\bar{\eta}_x^i) - h(\bar{\eta})\right),
\end{aligned} \tag{2.1.7}$$

where η_x^i agrees with η except at x, where it takes the value i. As before,

$$\left(f(\bar{\eta}_x^i) - f(\bar{\eta})\right) \left(h(\bar{\eta}_x^i) - h(\bar{\eta})\right) \geq 0,$$

whenever f and h are increasing. Therefore, (2.1.3) holds for the coupled process.

Next, one would like to know that the coupled process is also attractive. For $\bar{\eta} \leq \bar{\vartheta}$, the rates of the coupled process satisfy

$$
\begin{aligned}
c_0(x,\bar{\eta}) &\geq c_0(x,\bar{\vartheta}) && \text{if } \bar{\eta}(x) \neq 0,\\
c_0(x,\bar{\eta}) - c_0(x,\bar{\vartheta}) + c_1(x,\bar{\eta}) &\geq c_1(x,\bar{\vartheta}) && \text{if } \bar{\eta}(x) = 2,\\
c_1(x,\bar{\vartheta}) + c_2(x,\bar{\vartheta}) - c_2(x,\bar{\eta}) &\geq c_1(x,\bar{\eta}) && \text{if } \bar{\eta}(x) = 0,\\
c_2(x,\bar{\vartheta}) &\geq c_2(x,\bar{\eta}) && \text{if } \bar{\eta}(x) \neq 2.
\end{aligned}
$$

In a similar manner as for the original spin system, one can use these inequalities to define a coupling $(\bar{\eta}_t, \bar{\vartheta}_t)$ such that

$$P^{(\bar{\eta},\bar{\vartheta})}(\bar{\eta}_t \leq \bar{\vartheta}_t) = 1, \tag{2.1.8}$$

for all time $t \geq 0$ whenever $\bar{\eta} \leq \bar{\vartheta}$. This has the consequence that the coupled process on $\{(\eta,\zeta) \in X^2 : \eta \leq \zeta\}$ is attractive. Therefore, by Theorem 2.1.1 the evolution of the coupled process also preserves positive correlations on this space.

One final note regarding Harris' Theorem and the coupled process. Suppose that death at the origin is suppressed in the smaller of the two processes. Then clearly one cannot expect to find a coupling such that $\eta_t \leq \zeta_t$ for all $t \geq 0$. However, one can maintain this inequality until the stopping time $R = \inf\{t : O \notin \zeta_t\}$. In particular, a modified coupling (η_t, ζ_t) is defined whereby for $t < R$, the rates for the coupled process are given by (2.1.6) and for $t \geq R$, the two process evolve independently. Thus,

$$\eta_t \subseteq \zeta_t \qquad \text{for all } t < R,$$

and there is no particular ordering relation that is guaranteed to hold after time R. It turns out that this modified coupled process is also attractive and satisfies (2.1.3). The fact that this modified coupled process satisfies (2.1.3) follows from the same argument as for the original coupled process. Attractiveness is not much harder to verify. Obey the rates that give rise to (2.1.8) until time R_1, which is the R corresponding to the smaller of the two coupled processes. In the interval $[R_1, R_2)$, three (nonindependent) copies of coupling (2.1.6) are used. Here R_2 is the R corresponding to the larger of the two coupled processes. The primary copy applies to the second of the two coupled processes. Then the first coordinate in the smaller of the two coupled processes is coupled to first coordinate in the larger of the coupled processes via coupling (2.1.6). Similarly, the second coordinates of each coupled process are coupled using (2.1.6). After time R_2, two independent copies of (2.1.6) continue to preserve the appropriate relations between the first coordinates of each coupled process and between the second coordinates of each coupled process respectively. Therefore, this modified process (η_t, ζ_t) in which $O \in \eta_t$ for all $t \geq 0$ satisfies the hypotheses of Harris' Theorem.

2.2 Characterizations of the Survival Phases

In this section, we study the model when $P^O(\eta_t \neq \emptyset \ \forall \ t) > 0$. We begin by asking whether or not weak survival can occur above the local survival threshold.

Salzano and Schonmann [27] proved that weak survival does not occur for the contact process on homogeneous graphs in the local survival phase. The properties of the contact process that their proof uses are that it is translation invariant, strong Markov, and attractive. Therefore, the probability of weak survival is zero above the local survival threshold for any translation invariant, attractive, strong Markov process on a homogeneous graph G taking values in $\{0,1\}^G$. In particular, when $P(\eta_t^A \neq \emptyset \ \forall \ t) > 0$,

$$P(\eta_t^A \neq \emptyset \ \forall \ t) = P(O \in \eta_t^A \text{ for unbounded } t), \tag{2.2.1}$$

for any finite initial configuration A.

Here is the main idea behind their proof. Let X_t be an attractive, strong Markov process taking values in $\{0,1\}^G$. They make the observation that local survival is almost surely equivalent to the event that for every $n \in \mathbb{N}$ there exists a finite time T_n such that the process contains a (fully occupied) ball of radius n centered at the origin. Using this fact, they prove that $P(O \in X_t^A \text{ for unbounded } t) > 0$, implies that

$$\lim_{n \to \infty} P(O \in X_t^{B(O,n)} \text{ for unbounded } t) = 1, \tag{2.2.2}$$

where $B(O,n)$ denotes the ball of radius n centered at the origin. On the event that the process survives, a ball of size n must become occupied somewhere. By the strong Markov property, the process can be restarted at this random time. Homogeneity of the graph and equation (2.2.2) imply that the probability of weak survival tends to zero as n tends to infinity. For a complete proof, see Salzano and Schonmann [27] Theorem 2(i) .

Next we turn our attention to the supercritical uniform model. Here the process converges in distribution to a measure that is a nontrivial convex combination of $\delta_\emptyset$ and $\delta_{\mathbb{T}^d}$. In fact,

$$P^A(\eta_t \in \cdot) \to P^A(\eta_t \neq \emptyset \ \forall\ t)\delta_{\mathbb{T}^d} + P^A(\eta_t = \emptyset \text{ some } t)\delta_\emptyset, \qquad (2.2.3)$$

for all finite, connected configurations A. This behavior is known as complete convergence. When $P^A(\eta_t \neq \emptyset \ \forall\ t) = 0$, it is immediate that complete convergence holds. It turns out that survival together with complete convergence is equivalent to supercriticality. In particular, complete convergence fails in the interval $(\beta_2(d), \beta_4(d))$. This equivalence has the interesting consequence that survival together with complete convergence is a monotone increasing property of the parameter β, a fact that is not apparent from the definition of complete convergence.

Salzano and Schonmann [27] also investigated the question of monotonicity of the complete convergence property for the contact process on general graphs. They determined that homogeneity of the graph, attractiveness of the process, and self duality could be used to prove the desired result. The definition of supercriticality for the uniform model is actually a modified version of the criterion given in Theorem 2(b) of [27] for complete convergence of the contact process. Here, the ideas used to prove Theorem 2(b) in [27] are adapted to prove the equivalence of supercriticality and complete convergence for the uniform model.

Lemma 2.2.1 *If $P(\eta_t^O \neq \emptyset\ \forall\ t) > 0$, then complete convergence holds if and only if*

$$\lim_{n\to\infty} \liminf_{t\to\infty} P(O \in \eta_t^{B(O,n)}) = 1.$$

In particular, if $P(\eta_t^O \neq \emptyset\ \forall\ t) > 0$ and complete convergence holds at β^*, then the same is true for all $\beta > \beta^*$.

Proof. First assume that $P(\eta_t^O \neq \emptyset\ \forall\ t) > 0$ and that complete convergence holds. Then $\lim_{t\to\infty} P(O \in \eta_t^{B(O,n)}) = P(\eta_t^{B(O,n)} \neq \emptyset\ \forall\ t)$. This together with equations (2.2.1) and (2.2.2) gives the if direction of the implication.

Assuming that $\lim_{n\to\infty} \liminf_{t\to\infty} P(O \in \eta_t^{B(O,n)}) = 1$, it is immediate that $P(\eta_t^O \neq \emptyset\ \forall\ t) > 0$. Given finite $A \subset \mathbb{T}^d$, let $T_n = \inf\{t : B(O,n) \subseteq \eta_t^A\}$. For $s < t$,

$$\begin{aligned} P(O \in \eta_t^A) &\geq P(O \in \eta_t^A \mid T_n \leq s) P(T_n \leq s) \\ &\geq \inf_{t-s \leq u} P(O \in \eta_u^{B(O,n)}) P(T_n \leq s), \end{aligned}$$

where the final inequality follows from the strong Markov property and attractiveness. Therefore, for all $s \in \mathbb{R}_+$ and $n \in \mathbb{N}$,

$$\liminf_{t\to\infty} P(O \in \eta_t^A) \geq \liminf_{t\to\infty} P(O \in \eta_t^{B(O,n)}) P(T_n \leq s),$$

Recall the observation that was made in [27]:

$$\lim_{n\to\infty} \lim_{s\to\infty} P(T_n \leq s) = P(O \in \eta_t^A \text{ for unbounded } t).$$

This together with equation (2.2.1) implies that

$$\liminf_{t\to\infty} P(O \in \eta_t^A) \geq P(\eta_t^A \neq \emptyset\ \forall\ t).$$

Since $P(O \in \eta_t^A) \leq P(\eta_s^A \neq \emptyset\ \forall\ s \leq t)$, it follows that $\limsup_{t\to\infty} P(O \in \eta_t^A) \leq P(\eta_t^A \neq \emptyset\ \forall\ t)$. Thus,

$$\lim_{t\to\infty} P(O \in \eta_t^A \mid \eta_t^A \neq \emptyset\ \forall\ t) = 1.$$

It follows that for all finite $B \subset \mathbb{T}^d$ $\lim_{t\to\infty} P(B \subseteq \eta_t^A \mid \eta_t^A \neq \emptyset\ \forall\ t) = 1$, which completes the proof. ∎

2.3 Absence of a Weak Survival Phase

As a consequence of connectedness and attractiveness, it turns out that $\beta_2(d) = \beta_3(d)$. Hence, the uniform model does not have an intermediate phase that is characterized by weak survival. Combining this with the fact that weak survival cannot happen above the local survival threshold, if the process survives, then it survives locally.

Proof of Theorem 1.4.1a). It suffices to show that $P(\eta_t^O \neq \emptyset\ \forall\ t) > 0$ implies that $P(O \in \eta_t^O \text{ for unbounded } t) > 0$. Let $\mathbb{B}_i^d = \{x \in \mathbb{T}^d : \|x - x_i\| \leq \|O - x\|\} \cup O$ where $x_1, \ldots, x_{d+1}$ denote the $d+1$ nearest neighbors of the root O. By rotational symmetry,

$$P(\mathbb{B}_i^d \cap \eta_t^O \neq \emptyset) \geq \frac{P(\eta_t^O \neq \emptyset)}{d+1} \geq \frac{P(\eta_s^O \neq \emptyset\ \forall\ s)}{d+1}.$$

Using the fact that the uniform model is an attractive spin system and that δ_O is positively correlated,

$$P(\mathbb{B}_i^d \cap \eta_t^O \neq \emptyset, \mathbb{B}_j^d \cap \eta_t^O \neq \emptyset) \geq P(\mathbb{B}_i^d \cap \eta_t^O \neq \emptyset) P(\mathbb{B}_j^d \cap \eta_t^O \neq \emptyset) \qquad i \neq j$$

by Theorem 2.1.1. Since η_t^O is connected, $P(O \in \eta_t^O) \geq P(\mathbb{B}_i^d \cap \eta_t^O \neq \emptyset, \mathbb{B}_j^d \cap \eta_t^O \neq \emptyset)$. Therefore, the assumption that $P(\eta_s^O \neq \emptyset \ \forall \ s) > 0$ implies that $P(O \in \eta_t^O)$ is bounded away from zero. Hence, $P(O \in \eta_s^O$ for unbounded $s) > 0$. ∎

Remark. A slight modification of this proof works for the double parameter uniform model. There connectedness of the single parameter model is replaced by connectedness of $\mathcal{G}(\eta)$.

2.4 The Rooted Chain

In order to obtain actual estimates on the critical values, it will be convenient to analyze the behavior of the uniform model on a single branch $\mathbb{B}^d$ of $\mathbb{T}^d$. Recall that $\mathbb{B}_i^d = \{x \in \mathbb{T}^d : \|x - x_i\| \leq \|O - x\|\} \cup O$ where $x_1, \ldots, x_{d+1}$ denote the $d+1$ nearest neighbors of the root O. Take $\mathbb{B}^d = \mathbb{B}_1^d$ and consider the initial configuration $\eta_0 = (\mathbb{T}^d \setminus \mathbb{B}^d) \cup O$. By connectedness, $\eta_t \supseteq \eta_0$ for all $t \geq 0$. Therefore, it suffices to keep track of the intersection with $\mathbb{B}^d$, namely $A_t = \eta_t \cap \mathbb{B}^d$. First, note that A_t is a connected subset of $\mathbb{B}^d$ since $\mathbb{B}^d$ and η_t are both connected and there is a unique path connecting any two vertices in $\mathbb{B}^d$. Also, $O \in A_t$ for all $t \geq 0$ since $O \in \eta_0$. Furthermore, $|A_t|$ is finite for all $t \geq 0$, where $|A| = |\{x : x \in$

$A \setminus O\}|$. To see this consider the rate at which $|A| \to |A| + 1$. This is given by $\beta\,|\{x \in A^c : \|x - A\| = 1\}|$, where $A^c = \mathbb{B}^d \setminus A$.

Proposition 2.4.1 *For all finite, connected $A \subset \mathbb{B}^d$ containing O,*

$$|\{x \in A^c : \|x - A\| = 1\}| = (d-1)|A| + 1. \tag{2.4.1}$$

Proof. If $|A| = 0$, then $A = \{O\}$ and $|\{x \in A^c : \|x - A\| = 1\}| = 1$. Assume that (2.4.1) holds for all $|A| < n$. Given $|A| = n$, choose $x \in A$ such that $A \setminus x$ is connected and contains O. By induction,

$$|\{y \in (A \setminus x)^c : \|y - A \setminus x\| = 1\}| = (d-1)(|A| - 1) + 1,$$

Adding x back into the set deletes one element from

$$\{y \in (A \setminus x)^c : \|y - A \setminus x\| = 1\},$$

namely x, and adds the d neighbors of x to this set. Therefore,

$$|\{y \in A^c : \|y - A\| = 1\}| = |\{y \in (A \setminus x)^c : \|y - A \setminus x\| = 1\}| - 1 + d,$$

which proves the assertion. ∎

Let Y_t be a pure birth process such that

$$n \to n+1 \qquad \text{at rate} \qquad \beta\left((d-1)n + 1\right).$$

As a consequence of Proposition 2.4.1, we can couple A_t and Y_t such that

$$|A_t| \leq Y_t \qquad \text{for all } t \geq 0.$$

Let τ_n be the length of time that Y_t spends in state n. Then τ_n is exponential with mean $1/\beta\left((d-1)n+1\right)$. Since

$$\mathbb{E}\left(\sum_{n=0}^{\infty} \tau_n\right) = \infty,$$

it follows that $\sum_{n=1}^{\infty} \tau_n = \infty$ almost surely. Therefore, $Y_t < \infty$ almost surely for all $t \geq 0$ and consequently, $|A_t| < \infty$ almost surely for all $t \geq 0$.

The Markov chain A_t will be referred to as the *rooted chain.* As noted above, it is irreducible with state space $\mathcal{C}_d = \{\text{finite, connected } A \subset \mathbb{B}^d \text{ containing } O\}$ and rates

$$q(A,B) = \begin{cases} c(x, \eta_0 \cup A) & \text{if } B = A \cup x \text{ or } B = A \setminus x, \\ 0 & \text{otherwise,} \end{cases}$$

for $A,\ B \in \mathcal{C}_d$. Let $\pi(A) = \beta^{|A|}$, where we have made the convention that the cardinality of A is the number of vertices in $A \setminus O$. Since

$$\pi(A)q(A, A \cup \{x\}) = \beta^{|A|+1} = \pi(A \cup \{x\})q(A \cup \{x\}, A)$$

for all $x \in \mathbb{B}^d$ such that $\|x - A\| = 1$, the rooted chain is reversible with respect to the measure $\pi(\cdot)$. For $A \in \mathcal{C}_d$, say that $x \in A$ is a leaf if $x \neq O$ and $|\{y \in A : \|x - y\| = 1\}| = 1$. Denote the set of all vertices in A that are leaves by ∂A. The connection between the behavior of the finite interacting particle system and the the rooted chain is outlined in the next theorem.

Theorem 2.4.2

a) If the rooted chain is positive recurrent, then the uniform model is subcritical, i.e. $\mathbb{E}^O(\tau) < \infty$.

b) If the rooted chain is transient, then uniform model is supercritical, i.e.

$$\lim_{n\to\infty} \liminf_{t\to\infty} P^{B(O,n)}(O \in \eta_t) = 1.$$

Proof. Let ξ_t denote the product of $d+1$ independent copies of the rooted chain with initial state $\{O\}$. Paste together the $d+1$ roots, one on top of the other, and locate the roots at the origin of $\mathbb{T}^d$. By this correspondence, the product chain is equal in distribution to a uniform model on $\mathbb{T}^d$ with death at O suppressed. Let η_t^O denote the uniform model on $\mathbb{T}^d$ with initial state O. By attractiveness, we can couple η_t^O and ξ_t such that

$$\eta_t^O \subseteq \xi_t \qquad \forall\, t \geq 0. \tag{2.4.2}$$

Furthermore, for any initial configuration A containing O we can couple η_t^A and ξ_t such that

$$\xi_t \subseteq \eta_t^A \qquad \forall\, 0 \leq t < R, \tag{2.4.3}$$

where $R = \inf\{t : O \notin \eta_t^A\}$ (see the final paragraph of Section 2.1).

The positive recurrence of the rooted chain is equivalent to positive recurrence of the product chain. This follows from the fact that the reversible measure for the product chain is the product of $d+1$ copies of the reversible measure for the rooted chain, which is certainly summable. Let $T_0 = 0$. For $i \geq 1$, set $T'_{i-1} = \inf\{t > T_{i-1} : \xi_t \neq \{O\}\}$ and $T_i = \inf\{t > T'_{i-1} : \xi_t = \{O\}\}$. Thus T_i denotes the time at which the product chain makes its ith visit to $\{O\}$. Let $N = \min\{n : \eta^O_{T_n} = \emptyset\}$. By the strong Markov property and containment (2.4.2), N is geometric with parameter $p = P(\eta^O_{T_1} = \emptyset)$. Therefore,

$$\mathbb{E}^O(\tau) \leq \mathbb{E}(T_N) = \mathbb{E}\left(\sum_{i=1}^{N}(T_i - T_{i-1})\right) = \mathbb{E}(N)\,\mathbb{E}(T_1)\,, \tag{2.4.4}$$

where the final equality is an application of Wald's Lemma. Positive recurrence of the product chain implies that $\mathbb{E}(T_1) < \infty$. Therefore, a) follows from (2.4.4).

Assume that the rooted chain is transient. As before, x_i $i = 1, \ldots, d+1$ denote the $d+1$ nearest neighbors of the origin. Let

$$S = \inf\{s : \xi_t \supseteq \{O, x_1, \ldots, x_{d+1}\} \text{ for all } t \geq s\}.$$

Since the rooted chain is transient, $P(S < \infty) = 1$ and $P(S \leq u) > 0$ for all $u > 0$. By containment (2.4.3),

$$\begin{aligned}
P(O \in \eta^A_t \text{ for all } t \geq 0) &= P(O \in \eta^A_t \text{ for all } t \leq S) \\
&\geq P(O \in \eta^A_t \text{ for all } t \leq S \mid S \leq u)P(S \leq u) \\
&= P(O \in \eta^A_t \text{ for all } t \leq u \mid S \leq u)P(S \leq u),
\end{aligned}$$

for any initial configuration A containing O and $u > 0$. By the remarks in the final paragraph of Section 2.1, (η_t, ξ_t) is also attractive and satisfies condition (2.1.3) of Harris' Theorem. Since $\{O \in \eta_t^A \text{ for all } t \leq u\}$ and $\{S \leq u\}$ are increasing events and $\delta_A \times \delta_O$ is positively correlated, it follows, by Corollary 2.1.2, that

$$P(O \in \eta_t^A \text{ for all } t \leq u \mid S \leq u) \geq P(O \in \eta_t^A \text{ for all } t \leq u).$$

Thus,

$$P(O \in \eta_t^A \text{ for all } t \geq 0) \geq P(O \in \eta_t^A \text{ for all } t \leq u)P(S \leq u). \qquad (2.4.5)$$

By bound (2.4.5) and Lemma 2.2.1, it suffices to show that

$$\lim_{u\to\infty} \lim_{n\to\infty} P(O \in \eta_t^{B(O,n)} \text{ for all } t \leq u)P(S \leq u) = 1$$

If the origin becomes vacant at some time $t \leq u$, then there exists a time $s \leq u$ such that $\eta_s^{B(O,n)} \cap \mathbb{B}_i^d = \{O\}$ for at least d indices. Since $\frac{d^n-1}{d-1}$ is the number of vertices in $B(O,n) \cap \mathbb{B}^d \setminus \{O\}$,

$$P(\exists\ s \leq u \ni \eta_s^{B(O,n)} \cap \mathbb{B}^d = \{O\}) \leq (1 - e^{-u})^{\frac{d^n-1}{d-1}}.$$

It follows that $P(\exists\ s \leq u \ni \eta_s^{B(O,n)} \cap \mathbb{B}_i^d = \{O\}$ for at least d indices) tends to zero as n tends to infinity. Therefore,

$$\lim_{n\to\infty} P(O \in \eta_t^{B(O,n)} \text{ for all } t \leq u)P(S \leq u) = P(S \leq u).$$

Letting u tend to infinity completes the proof. ∎

Remark. The positive recurrence of the rooted chain is in fact equivalent to finite expected extinction time of the uniform model. In order to prove this, one would construct the shape chain for the uniform model, a Markov chain on the finite subsets of $\mathbb{T}^d$ where isomorphic sets are identified and that has a transition from the empty set to the singleton at rate β (see Section 4.2 for the details of the construction). The following string of equivalences proves the assertion: positive recurrence of the rooted chain is equivalent to positive recurrence of the product chain, which is equivalent to positive recurrence of the shape chain, which is equivalent to finite expected extinction time. The only statement that needs proof is the equivalence of positive recurrence of the product chain and the shape chain. Given the construction of the shape chain, verifying that the reversible measure of the shape chain is summable if and only if the reversible measure of the product chain is summable proves the assertion.

By Theorem 2.4.2, the behavior of the uniform model is determined except for the values of β for which rooted chain is null recurrent. A priori, one might be inclined to think that null recurrence can only happen for at most one value of the parameter β. For instance, translation invariant random walks on $\mathbb{Z}$ with finite mean step size are null recurrent if and only if the mean step size is zero. Additionally, one might be tempted to think that there are general theorems that have been, or can be, proved in this regard. The next example is designed to show that some chains do in fact exhibit an entire interval of null recurrent behavior.

Example. Consider a collection of birth and death chains indexed by a parameter λ. The rates are given by

$$q_\lambda(k, k+1) = \lambda\ b(k) \qquad \text{and} \qquad q_\lambda(k, k-1) = d(k) \qquad \text{for } k \in \mathbb{N}.$$

Here, both b and d are positive functions with domain $\mathbb{N}$, except that $d(0) = 0$. This chain is positive recurrent if and only if

$$\sum_{k=1}^{\infty} \prod_{j=0}^{k-1} \frac{b(j)}{d(j+1)} \lambda^k < \infty,$$

and this chain is transient if and only if

$$\sum_{k=1}^{\infty} \frac{1}{b(k) \prod_{j=0}^{k-1} \frac{b(j)}{d(j+1)} \lambda^{k+1}} < \infty.$$

Therefore, null recurrence is equivalent to both series diverging. A sufficient condition for both series to diverge is that

$$\frac{1}{\limsup_{k\to\infty} \left(\prod_{j=0}^{k-1} \frac{b(j)}{d(j+1)} \right)^{1/k}} < \lambda < \frac{1}{\liminf_{k\to\infty} \left(b(k) \prod_{j=0}^{k-1} \frac{b(j)}{d(j+1)} \right)^{1/k}}.$$

Therefore, in order to produce an interval of null recurrent behavior it suffices to choose b and d such that that the liminf and the limsup are different. Take $b \equiv 1$. Fix $0 < \delta < \alpha$ and set

$$d(k) = \begin{cases} 0 & \text{if } k = 0, \\ \alpha^{k-1}/\delta^k & \text{if } k \geq 2 \text{ and } k \text{ is even}, \\ \delta^{k-1}/\alpha^k & \text{if } k \text{ is odd}. \end{cases}$$

So, for $\lambda \in (\delta, \alpha)$, this chain is null recurrent.

This example raises a question as to when such intervals can be ruled out. The natural context in which to ask this question seems to be attractive Markov chains on partially ordered spaces for which the one step transitions are between comparable states, the context of Harris' Theorem. In this setting, all of the rates for increasing transition can be scaled by a parameter so that the model is stochastically increasing in this parameter. The issue is to determine under what conditions on the state space and the rates does null recurrent behavior happen for at most one of value of the parameter. The forthcoming analysis for the rooted chain will demonstrate that even for specific examples this question can be challenging.

2.5 The Finite Expected Extinction Time Threshold

By Theorem 2.4.2 and the remark following its proof, the positive recurrence threshold for the rooted chain agrees with $\beta_1(d)$. In this section, we compute the positive recurrence threshold for the rooted chain and thereby compute $\beta_1(d)$. In fact, two proofs of Theorem 1.4.1b) are given. The first takes advantage of generating function arguments, while the second is combinatorial in nature.

Proof of Theorem 1.4.1b). It suffices to show that the rooted chain is positive recurrent if and only if $\beta \leq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$. Since the rooted chain is reversible with respect to the measure $\pi(\cdot)$, positive recurrence is equivalent to the summability of the series

$$C(\beta) = \sum_{n=0}^{\infty} c(n)\beta^n, \tag{2.5.1}$$

where $c(n)$ is the number of $A \in \mathcal{C}_d$ such that $|A| = n$. The unique set of cardinality zero is $\{O\}$ so that $c_0 = 1$. For $n \geq 1$, the following recursion holds:

$$c(n) = \sum_{(k_1,\ldots,k_d)} c(k_1)\ldots c(k_d), \tag{2.5.2}$$

where the sum is taken over all d-tuples in $\mathbb{N}^d$ such that $k_1+\cdots+k_d = n-1$. To see this, note that $n \geq 1$ implies that x_1, the nearest neighbor of the root O, is in the set; otherwise, the set would be disconnected from O. Given that both O and x_1 are in the set, there are $n-1$ additional vertices in the set. Regarding x_1 as the root of d distinct copies of $\mathbb{B}^d$, choose $(k_1,\ldots,k_d)$ in $\mathbb{N}^d$ such that $\sum_{i=1}^d k_i = n-1$ and place k_i+1 vertices (including x_1) on the ith copy of $\mathbb{B}^d$. The number of distinct arrangements of k_i+1 vertices on $\mathbb{B}^d$ is $c(k_i)$, which proves recursion (2.5.2).

Multiplying recursion (2.5.2) by β^{n-1} and taking the sum from $n = 1$ to ∞ gives

$$\frac{C(\beta)-1}{\beta} = [C(\beta)]^d. \tag{2.5.3}$$

Let $p(y,\beta) = \beta y^d - y + 1$. If $C(\beta) < \infty$, then $p(C(\beta),\beta) = 0$. For each $\beta > 0$, $p(\cdot,\beta)$ is a strictly convex function on $\mathbb{R}^+$ with a unique minimum at $(\beta d)^{1-d}$. There exists a $y \in \mathbb{R}^+$ such that $p(y,\beta) = 0$ if and only if $p((\beta d)^{1-d},\beta) \leq 0$. Furthermore, $p((\beta d)^{1-d},\beta) \leq 0$ if and only if $\beta \leq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$, establishing the only if part.

Multiplying recursion (2.5.2) by β^{n-1} and taking the sum from $n = 1$ to N gives

$$\frac{C_N(\beta) - 1}{\beta} < [C_N(\beta)]^d, \tag{2.5.4}$$

where $C_N(\beta)$ denotes the partial sum to the Nth term of the series (2.5.1). Assume that $\beta \leq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$ and let $y_1(\beta) \leq y_2(\beta)$ denote the two positive roots of $p(\cdot, \beta)$. By inequality (2.5.4), $p(C_N(\beta), \beta) > 0$. Therefore, $C_N(\beta) \in (0, y_1(\beta)) \cup (y_2(\beta), \infty)$. At $\beta = \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$, $y_2(\beta) > 1$. As β decreases to 0, $y_2(\beta)$ increases to infinity, while $C_N(\beta)$ tends to 1. Hence, the statement that $C_N(\beta) \in (y_2(\beta), \infty)$ for some $\beta \leq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$ contradicts the continuity of $C_N(\beta)$ in β. Therefore, $C_N(\beta) \in (0, y_1(\beta))$ for all $\beta \leq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$ and for all $N \in \mathbb{N}$. Let N tend to infinity to obtain $C(\beta) \leq y_1(\beta) < \infty$. ■

It is well known in the Combinatorics literature that, in case $d = 2$, the unique solution to recursion (2.5.2) is the Catalan numbers, i.e.

$$c(n) = \frac{1}{n+1}\binom{2n}{n}. \tag{2.5.5}$$

Solving (2.5.3) for $C(\beta)$, gives

$$C(\beta) = \frac{1 - \sqrt{1 - 4\beta}}{2\beta}.$$

We choose the root with the negative sign since $\lim_{\beta \to 0} C(\beta) = 1$. Computing the power series for $C(\beta)$ centered at zero shows that $c(n)$ is in fact the nth Catalan number. By Stirling's formula,

$$c(n) \sim \frac{4^n}{\sqrt{2\pi} n^{3/2}}$$

which gives an alternate proof of summability up to and including 1/4 in case $d = 2$. Here, $\sim$ means that the ratio tends to one.

The technique used to compute $c(n)$ in case $d = 2$ becomes complicated and eventually breaks down. At $d = 5$, the Galois group is the entire symmetric group and therefore the roots are no longer computable by radicals. Therefore, one does not obtain an explicit expression for $C(\beta)$ from recursion (2.5.2). However, a simple combinatorial argument can be used to compute $c(n)$ for all $d \geq 2$. Consider the correspondence

$$\{A \in \mathcal{C}_d : |A| = n\} \leftrightarrow \{A \in \mathcal{C}_d : |A| = dn + 1 \text{ and } |\partial A| = (d-1)n + 1\}$$

that is given by mapping a set A of size n to the set B of size $dn + 1$ obtained by adding all vertices within distance one of A. The number of $A \in \mathcal{C}_d$ of cardinality $dn + 1$ with $(d-1)n + 1$ leaves is known to be $\binom{dn}{n} / ((d-1)n + 1)$: see Lemma 2.5.1 below. Therefore,

$$c(n) = \frac{1}{(d-1)n + 1} \binom{dn}{n}. \tag{2.5.6}$$

Again, an application of Stirling provides the desired summability. Verifying that (2.5.6) satisfies recursion (2.5.2) directly is not easy.

Lemma 2.5.1 *Let* $\mathcal{C}_d(n, \ell) = \{A : |A| = n \text{ and } |\partial A| = \ell\}$ *and let* $c(n, \ell) = |\mathcal{C}_d(n, \ell)|$. *For* $n \in \mathbb{N}$,

$$c(dn + 1, (d-1)n + 1) = \frac{1}{(d-1)n + 1} \binom{dn}{n}.$$

Proof. The first step in the proof is to show that there is a one-to-one correspondence between $A \in \mathcal{C}_d(dn+1,(d-1)n+1)$ and $(dn+1)$-tuples with n entries that are d and $(d-1)n+1$ entries that are zero such that $\sum_{j=1}^{k} y_j \geq k$ for $1 \leq k \leq dn$. Therefore, in order to prove the assertion, it will suffice to count the number of $(dn+1)$-tuples with these properties. This is the second step in the proof.

In order to prove the one-to-one correspondence, we first need to fix an ordering of the vertices of $\mathbb{B}^d$. Associate to each $x \in \mathbb{B}^d$ a level $\ell(x)$ that is given by $\|x - O\|$. Say that y is a child of x (or equivalently that x is the parent of y) if $\ell(y) = \ell(x)+1$ and $\|y - x\| = 1$. Without loss of generality, fix an ordering of the children of each vertex. This induces a total ordering on the vertices of $\mathbb{B}^d$ that is given by $x < y$ if one of the following holds:

1) $\ell(x) < \ell(y)$;

2) $\ell(x) = \ell(y)$, x and y are children of a common vertex z, and $x < y$; or

3) $\ell(x) = \ell(y)$ and the level of the parent of x is less than the level of the parent of y.

For $m \geq 1$ and $A \in \mathcal{C}_d(m) = \{A \in \mathcal{C}_d : |A| = m\}$, let $\{x_i\}_{i=1}^{m}$ be the vertices in $A \setminus O$ ordered so that $x_i < x_{i+1}$ for $i = 1, \ldots, m-1$. For $x \in A$, let $\tilde{c}(x) = \{y \in A : y \text{ is a child of } x\}$ be the children of x that are in A and let $y_i = |\tilde{c}(x_i)|$. Associate to A the m-tuple,

$$\Phi(A) = (y_1, \ldots, y_m).$$

In particular, $A \in \mathcal{C}_d(m)$ can be mapped into $\mathbb{N}^m$. First note that $\sum_{i=1}^m y_i$ is the number of edges in $A \setminus O$, which is $m-1$. A less obvious fact about this m-tuple is that $\sum_{i=1}^k y_i \geq k$ for $1 \leq k \leq m-1$. To see this, let $A_k = \{O, x_1\} \bigcup \{\cup_{i=1}^k \tilde{c}(x_i)\}$ and note that $A_k \in \mathcal{C}_d$ is a subset of A. If $\sum_{i=1}^k y_i = m-1$, then the assertion clearly holds. Otherwise, there exists $i > k$ such that $y_i > 0$, A_k is strictly contained in A, and more vertices must be added to A_k in order to obtain A. There are at most $1 + \sum_{i=1}^k y_i - k$ possible vertices to which these children may be added. This follows from the observation that $1 + \sum_{i=1}^k y_i$ is the total number of vertices in $A_k \setminus O$ and no more children of the vertices $\{x_1, \ldots, x_k\}$ can be added to the set. Therefore, $1 + \sum_{i=1}^k y_i - k \geq 1$ as desired. If in addition $m = dn+1$ for some $n \in \mathbb{N}$ and $|\partial A| = (d-1)n+1$, then $y_i = 0$ for exactly $(d-1)n+1$ indices. Since $\sum_{i=1}^{dn+1} y_i = dn$ and $y_i \leq d$, it must be the case that the remaining n indices satisfy $y_i = d$. Therefore, $\Phi(A)$ is a $(dn+1)$-tuple with n entries that are d and $(d-1)n+1$ entries that are zero such that $\sum_{i=1}^k y_i \geq k$ for $1 \leq k \leq dn$.

Let $(y_1, \ldots, y_{dn+1})$ be a $(dn+1)$-tuple with n entries that are d and $(d-1)n+1$ entries that are zero such that

$$\sum_{j=1}^k y_j \geq k \quad \text{for all} \quad 1 \leq k \leq dn. \tag{2.5.7}$$

We will construct a subtree $A \in \mathcal{C}_d(dn+1, (d-1)n+1)$ from this $(dn+1)$-tuple by constructing a sequence of subsets of vertices $\{O_k\}$. The set O_k will be called the open vertices at step k. The construction begins by setting $O_0 = \{x_1\}$, where x_1 denotes the nearest neighbor of the root. For $k \geq 1$, if $O_{k-1} \neq \emptyset$, then construct O_k from O_{k-1} by deleting the smallest element and adding the y_k children of the smallest element of O_{k-1}. Since $y_k = 0$ or d, it is not ambiguous which y_k children to add. In particular, $|O_k| = |O_{k-1}| + (y_k - 1)$. Iterating this relation and using the fact that $|O_0| = 1$, gives $|O_k| = 1 + \sum_{j=1}^{k} \Big(y_j - 1\Big)$. Assumption (2.5.7) implies that $|O_k| \geq 1$ for all $0 \leq k \leq dn$. The fact that $\sum_{i=1}^{dn+1} y_i = dn$ implies that $|O_{dn+1}| = 0$. Set

$$\Psi(y_1, \ldots, y_{dn+1}) = O \bigcup \{\cup_{k=0}^{dn+1} O_k\}.$$

To see that $\Psi(y_1, \ldots, y_{dn+1}) \in \mathcal{C}_d$, let $A_k = O \bigcup \{\cup_{i=0}^{k} O_i\}$. In particular, $A_{dn+1} = \Psi(y_1, \ldots, y_{dn+1})$. It is immediate that $A_0 \in \mathcal{C}_d$. Since A_{k+1} is A_k together with the y_{k+1} children of the smallest element of O_k and since $O_k \subset A_k$, it follows by induction that $A_{k+1} \in \mathcal{C}_d$. In order to verify that $|A_{dn+1}| = dn + 1$, observe that $|A_0| = 1$ and that $|A_{k+1}| = |A_k| + y_{k+1}$. Iterating this equation gives the result. After ordering the vertices of $\Psi(y_1, \ldots, y_{dn+1})$, the construction implies that $|\check{c}(x_i)| = y_i$. Therefore, $|\partial A_{dn+1}| = (d-1)n + 1$ and Φ is Ψ inverse.

We have shown that there is a one-to-one correspondence between $(dn + 1)$-tuples with n entries that are d and $(d-1)n + 1$ entries that are zero such that $\sum_{j=1}^{k} y_j \geq k$ for $1 \leq k \leq dn$ and $A \in \mathcal{C}_d(dn + 1, (d-1)n + 1)$. Therefore, in order to prove the assertion, it suffices to count the number of $(dn+1)$-tuples with these properties. This is accomplished by showing first that for any $(dn+1)$-tuple with n

entries that are d and $(d-1)n+1$ entries that are zero there exists a unique cyclic permutation of this $(dn+1)$-tuple that satisfies assumption (2.5.7). Given this, the number of $(dn+1)$-tuples with n entries that are d and $(d-1)n+1$ entries that are zero is $\binom{dn+1}{n}$. Given any such $(dn+1)$-tuple, there are exactly $dn+1$ distinct cyclic permutations of this $(dn+1)$-tuple; otherwise, two distinct permutations satisfy assumption (2.5.7). Therefore, the number of $(dn+1)$-tuples with n entries that are d and $(d-1)n+1$ entries that are zero that satisfy assumption (2.5.7) is $\binom{dn+1}{n}/(dn+1)$.

In order to show that this cyclic permutation exists, fix a $(dn+1)$-tuple with n entries that are d and $(d-1)n+1$ entries that are zero and assume that assumption (2.5.7) does not hold. Let k_1 be the minimum k such that $\sum_{j=1}^{k} y_j < k$. If $k_i < dn+1$, then proceed by letting k_{i+1} be the minimum $k > k_i$ such that $\sum_{j=k_i+1}^{k} y_j < k-k_i$. Repeat this procedure until no such k exists. This determines a finite collection of indices $\{k_1, \ldots, k_m\}$. Cyclically permute the $(dn+1)$-tuple so that y_{k_m+1} is the first entry:

$$(y_{k_m+1}, \ldots, y_{dn+1}, y_1, \ldots, y_{k_m}).$$

Since $\sum_{j=k_i+1}^{k} y_j \geq k - k_i$ for $k_i + 1 \leq k < k_{i+1}$, it follows that $\sum_{j=k_i+1}^{k_{i+1}} y_j = k_{i+1} - k_i - 1$. Therefore, $\sum_{j=1}^{k_m} y_j = k_m - m$. Using the fact that $\sum_{i=1}^{dn+1} y_i = dn$, $\sum_{j=k_m+1}^{dn+1} y_j = dn - k_m + m$. Thus, assumption (2.5.7) holds for the reordered $(dn+1)$-tuple. In particular, there is at least one cyclic permutation of this $(dn+1)$-tuple that satisfies assumption (2.5.7).

We now proceed to show that there is exactly one such permutation. Fix $(y_1, \ldots, y_{dn+1})$ with n entries that are d and $(d-1)n+1$ entries that are zero. Without loss of generality, assumption (2.5.7) holds. Consider

$$(y_{m+1}, \ldots, y_{dn+1}, y_1, \ldots, y_m) \quad 1 \le m \le dn.$$

By assumption, $\sum_{i=1}^{m} y_i \ge m$. Since $\sum_{i=1}^{dn+1} y_i = dn$, it follows that

$$\sum_{i=m+1}^{dn+1} y_i \le dn - m,$$

which is strictly less than the number of terms that appear in that summation. Consequently, assumption (2.5.7) does not hold for this permutation. ∎

2.6 The Complete Convergence Threshold: An Easy Bound

Since $\beta_1(d)$ is explicitly known, this gives a lower bound on the other critical values. Our attention now turns to obtaining upper bounds. If the total birth rate at a leaf is greater than the death rate, the boundary of the occupied set should have a net drift outward. Furthermore, it seems reasonable to expect this drift out at the boundary to force the occupied set to expand in all directions resulting in total occupation of the tree. We formalize this intuition and obtain an easy upper bound on $\beta_4(d)$.

Theorem 2.6.1 *For $d \ge 2$, $\beta_4(d) \le 1/d$.*

Proof. By Theorem 2.4.2, it suffices to show that the rooted chain is transient for $\beta > \frac{1}{d}$. Modify the rates $q(A,B)$ by suppressing all births at neighbors of nonleaves. To be precise, let $\mathcal{L}_d = \{A \in \mathcal{C}_d : A \text{ has exactly one leaf}\} \cup \{O\}$ and for $A, B \in \mathcal{C}_d$ define

$$\tilde{q}(A,B) = \begin{cases} q(A,B) & \text{if } A, B \in \mathcal{L}_d, \\ 0 & \text{otherwise.} \end{cases}$$

Let L_t denote the Markov chain with state space $\mathcal{L}_d$ and rates $\{\tilde{q}(A,B)\}$. If $A_1 \in \mathcal{L}_d$, $A_2 \in \mathcal{C}_d$, $A_1 \subseteq A_2$, $x \in A_1$, and $y \notin A_2$, then

$$\tilde{q}(A_1, A_1 \setminus x) \geq q(A_2, A_2 \setminus x) \qquad \text{and} \qquad \tilde{q}(A_1, A_1 \cup y) \leq q(A_2, A_2 \cup y).$$

Therefore, we can couple L_t and A_t such that $L_t \subseteq A_t$ for all $t \geq 0$. Consequently, if L_t is transient, then so is A_t. Since $|L_t|$ is a birth and death chain with birth rate $d\beta$ and death rate one, L_t is transient for $\beta > 1/d$. ∎

The positive recurrence and easy transience bounds of Theorems 1.4.1a) and 2.6.1 respectively are the analogs of the lower and upper bounds

$$\frac{1}{d\gamma}\left(\frac{d-1}{d(1-\gamma)}\right)^{d-1} - 1 \leq \lambda_2(d,\gamma) \leq \frac{1}{d}\left(\frac{1-\gamma d}{\gamma}\right) \tag{2.6.1}$$

given by Liggett [18] for the two parameter uniform model. To see this, multiply by γ and let γ decrease to zero. The technique used here to compute the positive recurrence threshold is almost the same as that used by Liggett to compute the lower bound for the double parameter uniform model. However, Liggett used a more sophisticated technique to obtain the upper bound that involved the Dirichlet principle and a notion that he called monotonicity. Essentially, he used these tools to restrict attention to the evolution of an embedded line process. Unfortunately, the simple coupling argument given here does not extend to the double parameter model.

2.7 The Complete Convergence Threshold: Improved Bounds via Flows

For reversible Markov chains, there is a very nice characterization of transience in terms of flows. A flow is a collection of real numbers corresponding to ordered pairs of states of the chain. It turns out that the existence of a flow with certain properties is equivalent to transience [23]. Therefore, in order to prove that β is in the supercritical phase, it suffices to exhibit a flow for the rooted chain with the desired properties.

The purpose of this section is to outline a method for constructing flows for the rooted chain that have these special properties. We begin with the definition of an antisymmetric, incompressible flow. Then, a method for constructing an entire class of such flows is described. Finally, the condition for transience in terms of flows is given. Thus the problem of transience is reduced to exploring particular instances of the construction and determining for which values β this condition holds.

Definition 2.7.1 *Given a Markov chain with state space S, an anti-symmetric, incompressible flow on S is a collection of real numbers, or weights, $\{w(x,y)\}$ indexed by $S \times S$ that satisfy the following properties:*

i) (Anti-Symmetry) For all $x, y \in S$, $w(x,y) = -w(y,x)$;

ii) (Existence of a Source) There exists a $x_0 \in S$

$$\sum_{y \in S} w(x_0, y) \neq 0; \tag{2.7.1}$$

iii) (Incompressibility) For $x \in S \setminus x_0$,

$$\sum_{y \in S} w(x,y) = 0. \tag{2.7.2}$$

For the rooted chain, the following construction leads to a class of flows on $\mathcal{C}_d$ that satisfy i) ii), and iii) above. Given a collection of weights $\{w(A,B)\}$, let

$$f(A) = \sum_{\{B : B \subseteq A\}} w(B,A) \qquad \text{for } A \neq \{O\} \tag{2.7.3}$$

be the *net flow into* A (from below). Given $A \in \mathcal{C}_d$, denote the neighbors of A that contain A by $\mathcal{N}_d(A)$. For $A \in \mathcal{C}_d$, say that $r(A,\cdot)$ is a *routing vector* if the support of $r(A,\cdot)$ is contained in $\mathcal{N}_d(A)$ and $\sum_{\{B : B \in \mathcal{N}_d(A)\}} r(A,B) = 1$. Note that $r(A,\cdot)$ is not required to be nonnegative. Given a collection of routing vectors, construct the flow recursively:

1) Set $f(\{O\}) = 1$.

2) If $f(A)$ is defined for all $|A| < n$, for each B such that $|B| = n$ set

$$w(A, B) = f(A)r(A, B) \tag{2.7.4}$$

for all A such that $|A| \leq n$, where it is understood that $f(A)r(A, B) = 0$ when $|A| = |B| = n$. Using equation (2.7.4), $f(B)$ is now defined by equation (2.7.3) for each B such that $|B| = n$. Thus, 2) may be repeated for all $|B| = n + 1$, $|A| \leq n + 1$.

3) For A, $B \in \mathcal{C}_d$ such that $|A| > |B|$, set $w(A, B) = -w(B, A)$.

Denote the collection $\{w(A, B)\}$ by F. Property 3) guarantees that F satisfies the anti-symmetry condition. By construction, $\sum_A w(\{O\}, A) = w(\{O\}, \{O, x_1\}) = 1$ so that F has a source. Take $B \neq \{O\}$ and combine equations (2.7.3) and (2.7.4) to obtain

$$\begin{aligned} \sum_{\{C \in \mathcal{C}_d : C \subseteq B\}} w(C, B) &= f(B) \\ &= f(B) \sum_{A \in \mathcal{N}_d(B)} r(B, A) \\ &= \sum_{A \in \mathcal{N}_d(B)} w(B, A). \end{aligned} \tag{2.7.5}$$

Since $w(B, A) = 0$ for all A such that $A \notin \mathcal{N}_d(B)$ and $B \notin \mathcal{N}_d(A)$, equation (2.7.5) proves incompressibility. This proves the next proposition.

Proposition 2.7.2 *Specifying a collection of routing vectors determines an anti-symmetric, incompressible flow.*

Theorem 2.7.3 ***(Lyons).*** *Given a continuous time irreducible, reversible Markov chain X_t with state space S, transition rates $q(x,y)$, and reversible measure π, transience of X_t is equivalent to the existence of an anti-symmetric, incompressible flow $\{w(x,y)\}$ on S such that*

$$\sum_{x,y\in S} \frac{w^2(x,y)}{\pi(x)q(x,y)} < \infty, \tag{2.7.6}$$

where, by convention, $0/0 = 0$ and $a/0 = \infty$ when $a > 0$.

The series given in condition (2.7.6) is known as the kinetic energy series, or simply the energy series. Thus, if condition 2.7.6 holds, the flow is said to have finite energy. We denote the energy by $\mathcal{K}(F)$. The existence of such a flow is equivalent to the existence of a nonconstant bounded function on the state space of the Markov chain that is harmonic, except at a single state x_0 where it attains its minimum value. The existence of such a function is a well known criterion for transience. The connections between these ideas will be more fully developed in Section 4.1. For now, we investigate some applications of this theorem to the rooted chain.

2.8 The Uniformly Routed Flow

With the general method of constructing flows on $\mathcal{C}_d$ outlined in Section 2.7, we attempt to a construct a flow that proves the transience of the rooted chain for $\beta > (d-1)^{d-1}/d^d$. Using the fact that $\pi(A)q(A,B) = \beta^{\max(|A|,|B|)}$, the kinetic energy series is

$$\mathcal{K}(F) = 2\sum_{n=0}^{\infty}\left(\frac{1}{\beta}\right)^{n+1}\sum_{A\in\mathcal{C}_d(n)}\sum_{B\in\mathcal{N}_d(A)} w^2(A,B).$$

Recall that $\mathcal{C}_d(n) = \{A \in \mathcal{C}_d : |A| = n\}$. Since β appears in the denominator, it is natural to try to maximize the radius of convergence by minimizing the coefficients. As a first attempt, fix $A \in \mathcal{C}_d(n)$ and

$$\text{minimize} \sum_{B\in\mathcal{N}_d(A)} w^2(A,B) \qquad \text{subject to} \qquad f(A) = \sum_{B\in\mathcal{N}_d(A)} w(A,B).$$

The solution to this minimization problem is to set

$$w(A,B) = \frac{f(A)}{|\mathcal{N}_d(A)|}.$$

By Proposition 2.4.1, $|\mathcal{N}_d(A)| = (d-1)|A| + 1$ so that

$$r(A,B) = \frac{1}{(d-1)|A|+1} \qquad \forall\ B \in \mathcal{N}_d(A). \tag{2.8.1}$$

In this case, the routing vectors are nonnegative. Let $h(n) = \sum_{A\in\mathcal{C}_d(n)} f^2(A)$. If $r(A,B)$ is defined by equation (2.8.1), then

$$\mathcal{K}(F) = 2\sum_{n=0}^{\infty}\frac{h(n)}{\beta^{n+1}((d-1)n+1)}.$$

By Theorems 2.7.3 and 2.4.2,

$$\beta_4(d) \le \limsup_{n\to\infty} h(n)^{1/n}. \tag{2.8.2}$$

Theorem 1.4.1c) will be a consequence of obtaining bounds on the limiting behavior of the sequence $h(n)^{1/n}$.

The first thing to note is that f can be computed exactly (see the next lemma). However, we will not be able to compute h explicitly. Instead, using the expression for f, h is expressed as a ratio. The goal is to prove that the sequence $h(n)$ is bounded above and below by sequences for which the associated power series have the same radius of convergence. Therefore, determining the radius of convergence of $\mathcal{K}(F)$ will be equivalent to determining the radius of convergence for a power series with coefficients equal to either the upper or lower bound. The bounds are chosen so that the numerators agree with the numerators of $h(n)$. The reason for choosing the bounds this way is to exploit the fact that the numerators of $h(n)$ satisfy a nice recursion. By choosing the denominator of the lower bound appropriately, the numerator recursion will guarantee that the lower bound satisfies a related recursion. The fact that the lower bound satisfies this related recursion allows one to obtain bounds on the radius of convergence of the power series with coefficients that agree with the lower bound.

We begin by finding an explicit expression for f. Then a combinatorial lemma is presented. As a consequence of this lemma, the numerator recursion for the sequence $h(n)$ is obtained. Next, the sequences that bound $h(n)$ are introduced. Finally, bounds on the radius of convergence of the power series with coefficients that agree with the lower bound are obtained for $d \geq 2$. This bound is an improvement over the easy transience bound of $1/d$ if and only if $d \geq 4$.

Definition 2.8.1 *An increasing path from $\{O\}$ to A in $\mathcal{C}_d$ is a collection $\{B_i\}_{i=0}^{|A|}$ of sets in $\mathcal{C}_d$ such that $B_0 = \{O\}$, $B_{|A|} = A$, and $B_{i+1} \in \mathcal{N}_d(B_i)$ for $i = 0, \ldots, |A|-1$. Let $N(A)$ be the number of paths that increase from $\{O\}$ to A.*

Lemma 2.8.2 *If $r(A, B)$ is defined by equation (2.8.1), then for A such that $|A| = n \geq 1$,*

$$f(A) = \frac{N(A)}{\prod_{k=0}^{n-1}((d-1)k+1)}.$$

Proof. If $|A| = 1$, then $A = \{O, x_1\}$. Since $r(\{O\}, \{O, x_1\}) = 1$, equation (2.7.3) gives $f(\{O, x_1\}) = 1$ as desired. Assume that the assertion holds for $|A| < n$. If $|A| = n$, then, by equations (2.7.3) and (2.7.4),

$$\begin{aligned} f(A) &= \sum_{\{B: B\subseteq A\}} w(B, A) = \sum_{\{B: B\subseteq A\}} f(B) r(B, A) \\ &= \sum_{\{B: A\in\mathcal{N}_d(B)\}} \frac{N(B)}{\prod_{k=0}^{n-2}((d-1)k+1)} \frac{1}{(d-1)(n-1)+1} \\ &= \frac{1}{\prod_{k=0}^{n-1}((d-1)k+1)} \sum_{\{B: A\in\mathcal{N}_d(B)\}} N(B) \\ &= \frac{N(A)}{\prod_{k=0}^{n-1}((d-1)k+1).} \qquad \blacksquare \end{aligned}$$

Lemma 2.8.3 *For $n \geq 1$, there exists a one-to-one correspondence between $\mathcal{C}_d(n)$ and the disjoint union $\sqcup_{(j_1,\ldots,j_d)} \mathcal{C}_d(j_1) \times \cdots \times \mathcal{C}_d(j_d)$, where the union runs over all*

d-tuples in $\mathbb{N}^d$ with $j_1 + \cdots + j_d = n-1$, such that under this correspondence

$$N(A) = \binom{|A|-1}{|A_1|, \ldots, |A_d|} N(A_1) \ldots N(A_d). \tag{2.8.3}$$

Proof. Let $\{y_1, \ldots, y_d\}$ denote the d nearest neighbors of x_1 in $\mathbb{B}^d$. Set $\mathbb{B}^d_{1i} = \{y : \|y_i - y\| \le \|x_1 - y\|\} \cup x_1$ and $A_i = A \cap \mathbb{B}^d_{1i}$. Since $\mathbb{B}^d_{1i} \cong \mathbb{B}^d$ for $1 \le i \le d$,

$$A \leftrightarrow (A_1, \ldots, A_d).$$

Equation (2.8.3) is an immediate consequence of this correspondence. ∎

Squaring equation (2.8.3) gives

$$N^2(A) = \binom{|A|-1}{|A_1|, \ldots, |A_d|}^2 N^2(A_1) \ldots N^2(A_d) \tag{2.8.4}$$

for $A \in \mathcal{C}_d$ such that $|A| \ge 1$. For $n \in \mathbb{N}$, let $N_n = \sum_{A \in \mathcal{C}_d(n)} N^2(A)$. Summing (2.8.4) over all ordered d-tuples $(A_1, \ldots, A_d) \in \mathcal{C}_d \times \cdots \times \mathcal{C}_d$ such that $|A_1| + \cdots + |A_d| = n-1$ gives

$$N_n = \sum_{(j_1, \ldots, j_d)} \binom{n-1}{j_1, \ldots, j_d}^2 N_{j_1} \ldots N_{j_d} \qquad \text{for } n \ge 1, \tag{2.8.5}$$

where the sum runs over all d-tuples in $\mathbb{N}^d$ with $j_1 + \cdots + j_d = n-1$. By definition of $h(n)$, Lemma 2.8.2, and definition of N_n,

$$h(n) = \frac{N_n}{\prod_{k=0}^{n-1} \left((d-1)k+1\right)^2}.$$

If we could solve recursion (2.8.5), then we would be able to compute $h(n)$ exactly. We pursue an alternate strategy and use recursion (2.8.5) to obtain information about the asymptotic behavior of $h(n)$. For $n \geq 1$,

$$(d-1)^{n-1}(n-1)! \leq \prod_{k=0}^{n-1}((d-1)k+1) \leq (d-1)^n n!.$$

Set $\ell(0) = 1$ and $u(0) = 1$. For $n \geq 1$, set

$$\ell(n) = \frac{N_n}{(d-1)^{2n} n!^2} \qquad \text{and} \qquad u(n) = \frac{N_n}{(d-1)^{2(n-1)}(n-1)!^2}.$$

Then

$$\ell(n) \leq h(n) \leq u(n) \quad \text{for all } n \in \mathbb{N}. \tag{2.8.6}$$

Furthermore, for $n \geq 1$, $u(n) = (d-1)^2 n^2 \ell(n)$ so that

$$\limsup_{n\to\infty} \ell(n)^{1/n} = \limsup_{n\to\infty} u(n)^{1/n}. \tag{2.8.7}$$

Combining inequality (2.8.6) and equation (2.8.7) proves the next proposition.

Proposition 2.8.4 $\limsup_{n\to\infty} h(n)^{1/n} = \limsup_{n\to\infty} \ell(n)^{1/n}$.

Proposition 2.8.5 *For* $n \geq 1$,

$$\ell(n) = \frac{1}{(d-1)^2 n^2} \sum_{(j_1,\ldots,j_d)} \ell(j_1)\cdots\ell(j_d), \tag{2.8.8}$$

where the sum runs over all d*-tuples in* $\mathbb{N}^d$ *with* $j_1 + \cdots + j_d = n-1$.

Proof. Divide recursion (2.8.5) by $(d-1)^{2n} n!^2$. ■

Finally, solving recursion (2.8.8) is equivalent to solving a modified recursion. Suppose that $\tilde{\ell}(0) = 1$ and for $n \geq 1$, $\tilde{\ell}(n)$ satisfies

$$\tilde{\ell}(n) = \frac{1}{n^2} \sum_{(j_1,\ldots,j_d)} \tilde{\ell}(j_1) \cdots \tilde{\ell}(j_d), \tag{2.8.9}$$

where the sum runs over all d-tuples in $\mathbb{N}^d$ with $j_1 + \cdots + j_d = n - 1$. Then $\ell(n) = \tilde{\ell}(n)/(d-1)^{2n}$ satisfies recursion (2.8.8). Therefore, obtaining bounds on the solution of recursion (2.8.9) gives bounds on the solution of recursion (2.8.8).

Theorem 2.8.6 *For $n \geq 1$,*

$$\tilde{\ell}(n) \leq \frac{1}{n}\left(\frac{d}{2}\right)^{n-1}.$$

Proof of Theorem 1.4.1c). By Theorem 2.8.6, the relationship between solutions of recursions (2.8.8) and (2.8.9), Proposition 2.8.4, and inequality (2.8.2), $\beta_4(d) \leq d/(2(d-1)^2)$. ∎

Theorem 2.8.6 is proved by induction. In order to execute the induction step, the following lemma is needed. This lemma is a special case of a well known expansion of the binomial coefficient $\binom{k+n-1}{n}$ with $k = 2$.

Lemma 2.8.7 *For any positive integer n,*

$$\sum_{j=1}^{n} \frac{2^j}{j!} \sum_{\gamma \in \Gamma(n,j)} \frac{1}{\gamma_1 \cdots \gamma_j} = n + 1, \tag{2.8.10}$$

where $\Gamma(n,j)$ *is the set of all ordered partitions of* n *of into* j *parts and* γ_i *is the* i*th element in the partition* γ*.*

Proof. Writing $-\log(1-x)$ as a power series centered at zero gives

$$(-\log(1-x))^j = \left(\sum_{k=1}^{\infty} \frac{x^k}{k}\right)^j = \sum_{n=j}^{\infty} \sum_{\gamma\in\Gamma(n,j)} \frac{1}{\gamma_1\cdots\gamma_j} x^n \qquad \text{for } |x|<1.$$

Let $k \in \mathbb{N}$ be such that $k \geq 1$. For $|x| < 1$,

$$\begin{aligned}
\sum_{n=1}^{\infty} \binom{k+n-1}{n} x^n &= \frac{1}{(1-x)^k} - 1 \\
&= e^{-k\log(1-x)} - 1 \\
&= \sum_{j=1}^{\infty} \frac{(-k\log(1-x))^j}{j!} \\
&= \sum_{j=1}^{\infty} \frac{k^j}{j!} \sum_{n=j}^{\infty} \sum_{\gamma\in\Gamma(n,j)} \frac{1}{\gamma_1\cdots\gamma_j} x^n \\
&= \sum_{n=1}^{\infty} \sum_{j=1}^{n} \frac{k^j}{j!} \sum_{\gamma\in\Gamma(n,j)} \frac{1}{\gamma_1\cdots\gamma_j} x^n.
\end{aligned}$$

Taking $k = 2$ completes the proof. ■

Proof of Theorem 2.8.6. By recursion (2.8.9), $\tilde{\ell}(1) = 1$ which verifies the assertion for $n = 1$. Assume that the assertion holds for $m < n$. We have

$$\begin{aligned}
\tilde{\ell}(n) &= \frac{1}{n^2} \sum_{(m_1,\ldots,m_d)} \tilde{\ell}(m_1)\cdots\tilde{\ell}(m_d) \\
&= \frac{1}{n^2} \sum_{j=1}^{\min(d,n-1)} \binom{d}{j} \sum_{\gamma\in\Gamma(n-1,j)} \tilde{\ell}(\gamma_1)\cdots\tilde{\ell}(\gamma_j),
\end{aligned}$$

since $\tilde{\ell}(m_i) = 1$ when $m_i = 0$. By assumption,

$$\begin{aligned}
\tilde{\ell}(n) &\le \frac{1}{n^2} \sum_{j=1}^{\min(d,n-1)} \binom{d}{j} \sum_{\gamma \in \Gamma(n-1,j)} \left(\frac{d}{2}\right)^{n-1-j} \frac{1}{\gamma_1 \cdots \gamma_j} \\
&\le \frac{1}{n^2} \sum_{j=1}^{\min(d,n-1)} \frac{d^j}{j!} \left(\frac{d}{2}\right)^{n-1-j} \sum_{\gamma \in \Gamma(n-1,j)} \frac{1}{\gamma_1 \cdots \gamma_j} \\
&= \frac{1}{n^2} \left(\frac{d}{2}\right)^{n-1} \sum_{j=1}^{\min(d,n-1)} \frac{2^j}{j!} \sum_{\gamma \in \Gamma(n-1,j)} \frac{1}{\gamma_1 \cdots \gamma_j} \\
&\le \frac{1}{n^2} \left(\frac{d}{2}\right)^{n-1} \sum_{j=1}^{n-1} \frac{2^j}{j!} \sum_{\gamma \in \Gamma(n-1,j)} \frac{1}{\gamma_1 \cdots \gamma_j}.
\end{aligned}$$

By Lemma 2.8.7,

$$\tilde{\ell}(n) \le \frac{1}{n} \left(\frac{d}{2}\right)^{n-1}. \qquad \blacksquare$$

A simple computation provides evidence that for large d the bound given in Theorem 1.4.1c) is close to the best that this flow achieves. Thus, not so much is lost in the inequality in Theorem 2.8.6. Let A_n be the discrete time Markov chain on $\mathcal{C}_d$ with transition probabilities defined by (2.8.1) and let ℓ_n be the number of leaves in the set A_n. By conditioning on ℓ_{n-1}, one gets a recursion that leads to

$$\mathbb{E}(\ell_n) = \frac{(d-1)n+1}{2d-1} \qquad n \ge 2.$$

In other words, the typical set that the uniform flow visits has a death rate that is roughly the birth rate divided by $(2d-1)\beta$. In $d = 2$, these sets are not only typical, but rather uniform flow visits them with very high probability:

$$\mathbb{E}\left(\ell_n^2\right) = \frac{(n+1)(5n+7)}{45} \qquad n \ge 4,\ d = 2,$$

and therefore,

$$\frac{\ell_n}{(d-1)n+1} \xrightarrow{P} \frac{1}{2d-1} \qquad \text{for } d = 2.$$

For small d the bound given in Theorem 1.4.1c) is much worse than $1/(2d-1)$. However, by handling the cases $d=2$ and $d=3$ separately, the bound induced on $\ell(n)$ by Theorem 2.8.6 can be improved to $14(n+1)(1/3)^{n+2}$ and $(1/5)^{n-1}$ respectively. We conjecture that $1/(2d-1)$ is the optimal bound for this flow. Numerical evidence suggests that one cannot hope for much better.

2.9 The Uniformly Distributed Flow

In the previous section, the main goal became to determine the asymptotic behavior of $h(n)$. This resulted from the fact that

$$\sum_{B \in \mathcal{N}_d(A)} r^2(A,B) = \frac{1}{(d-1)n+1},$$

and therefore, the presence of this factor did not affect the radius of convergence of $\mathcal{K}(F)$. If we require the routing vectors to be absolutely bounded by b, then

$$\frac{1}{(d-1)n+1} \leq \sum_{B \in \mathcal{N}_d(A)} r^2(A,B) \leq b^2((d-1)n+1). \tag{2.9.1}$$

Thus, under the assumption that routing vector are bounded, the asymptotic behavior of $h(n)$ governs the radius of convergence of $\mathcal{K}(F)$.

As a consequence of the construction, $\sum_{A \in \mathcal{C}_d(n)} f(A) = 1$. Hence, we seek to minimize a quadratic function subject to a linear constraint. If this linear constraint were the only constraint, then the solution would be to partition 1 into equal parts, i.e. distribute the fluid uniformly over sets of size n. However, we require the flow to be incompressible which introduces many additional constraints. Notice that if a flow exists with bounded routing vectors such that $f(A) = 1/c(|A|)$, then by inequality (2.9.1),

$$\mathcal{K}(F) \leq \frac{2b^2}{\beta} \sum_{n=0}^{\infty} \frac{(d-1)n+1}{c(n)\beta^n}. \tag{2.9.2}$$

This series is summable for $\beta > \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$ since, except for the factor of $(d-1)n+1$, each term is the exact reciprocal of the terms appearing in series (2.5.1). Due the these observations, we attempt to construct a uniformly distributed flow with bounded routing vectors.

Suppose that one has constructed routing vectors bounded by b such that $f(A) = 1/c(|A|)$ for all $A \in \mathcal{C}_d$ such that $|A| < n$. Exploit the fact that $\mathcal{C}_d(n)$ is in one-to-one correspondence with $\sqcup_{(k_1,\ldots,k_d)}\mathcal{C}_d(k_1) \times \cdots \times \mathcal{C}_d(k_d)$ where the union runs over all d-tuples in $\mathbb{N}^d$ such that $k_1+\cdots+k_d = n-1$ and use the routing vectors $\{r(A,\cdot)\}_{|A|<n}$ to construct the routing vectors for $\mathcal{C}_d(n)$. More specifically, associate to each set a *preliminary routing vector* $\alpha(A,i)$ that determines the amount of fluid routed to branch i in set A. In particular, let $\alpha(A,\cdot)$ be such that $\sum_{i=1}^{d} \alpha(A,i) = 1$. Again, $\alpha(A,i)$ is not required to be nonnegative. If A corresponds to $(A_1,\ldots,A_d)$, B corresponds to $(B_1,\ldots,B_d)$, $B \in \mathcal{N}_d(A)$, and $A_i \neq B_i$, then let

$$r(A,B) = \alpha(A,i)r(A_i,B_i).$$

Since

$$\sum_{\{B\in\mathcal{N}_d(A)\}} r(A,B) = \sum_{i=1}^{d} \sum_{\{B_i\in\mathcal{N}_d(A_i)\}} \alpha(A,i) r(A_i,B_i)$$
$$= \sum_{i=1}^{d} \alpha(A,i) \sum_{\{B_i\in\mathcal{N}_d(A_i)\}} r(A_i,B_i)$$
$$= \sum_{i=1}^{d} \alpha(A,i) = 1,$$

it follows that $r(A,\cdot)$ is a routing vector. Furthermore, if $|\alpha(A,i)| \leq 1$, then $r(A,B)$ is bounded by b. Therefore, in order to specify a collection of bounded routing vectors, it suffices to specify a collection $\alpha(A,i)$ of preliminary routing vectors that are absolutely bounded by one.

A priori, one might expect $\alpha(A,\cdot)$ to depend on the entire structure of A. However, it is reasonable to expect dependence only on the cardinalities of A_j for $1 \leq j \leq d$. One explanation for this is that the distribution that we are trying to achieve depends only on cardinality. A more practical reason for making this assumption is that it simplifies the set of equations that $\alpha(A,\cdot)$ must satisfy by allowing a second application of the induction hypothesis. For $k \in \mathbb{N}^d$ such that $k_1 + \cdots + k_d = n-1$, let $\{\alpha_i(n;k)\}_{i=1}^{d}$ be a preliminary routing vector in a set A when $|A| = n$ and $|A_j| = k_j$ for $1 \leq j \leq d$. Thus, the function $\alpha_i(n;k)$ must satisfy

$$\alpha_1(n;k) + \cdots + \alpha_d(n;k) = 1 \qquad \text{and} \qquad |\alpha_i(n;k)| \leq 1, \tag{2.9.3}$$

for all $n \geq 1$ and $k \in \mathbb{N}^d$ such that $k_1 + \cdots + k_d = n-1$. Also, require that for all permutations σ of d objects $\alpha_{\sigma(i)}(n;\sigma(k)) = \alpha_i(n;k)$ where σ acts on a d-vector in the usual manner by permuting the indices. This condition simply states that the preliminary routing vectors are invariant under automorphisms of $\mathbb{B}^d$. For all $A \in \mathcal{C}_d(n)$ and $B \in \mathcal{N}_d(A)$, set

$$r(A,B) = \alpha_i(n;k) r(A_i, B_i) \qquad \text{if } A_i \neq B_i, \tag{2.9.4}$$

where $|A| = n$ and $|A_j| = k_j$ for all $1 \le j \le d$. The goal is to choose $\alpha_i(n;\cdot)$ such that the flow is distributed uniformly over sets of size $n+1$.

For $B \in \mathcal{C}_d(n+1)$, set $k_i = |B_i|$. Make the convention that $c(-1) = 0$. Since $c(0) = 1$, $f(\{O\}) = 1/c(0)$ by definition. Proceeding inductively, the net flow into B is given by

$$\begin{aligned}
f(B) &= \sum_{\{A : B \in \mathcal{N}_d(A)\}} f(A) r(A,B) \\
&= \frac{1}{c(n)} \sum_{i=1}^{d} \sum_{\{A_i : B_i \in \mathcal{N}_d(A_i)\}} \alpha_i(n; k - e_i) r(A_i, B_i) \\
&= \frac{1}{c(n)} \sum_{i=1}^{d} \alpha_i(n; k - e_i) c(k_i - 1) \sum_{\{A_i : B_i \in \mathcal{N}_d(A_i)\}} f(A_i) r(A_i, B_i) \\
&= \frac{1}{c(n)} \sum_{i=1}^{d} \alpha_i(n; k - e_i) \frac{c(k_i - 1)}{c(k_i)},
\end{aligned} \tag{2.9.5}$$

where e_i is the d-vector with all entries equal 0 except the ith which is 1.

Lemma 2.9.1 *If, for each $n \ge 1$, there exists $\alpha_i(n;\cdot)$ satisfying (2.9.3) and*

$$\frac{c(n)}{c(n+1)} = \sum_{i=1}^{d} \alpha_i(n; k - e_i) \frac{c(k_i - 1)}{c(k_i)} \tag{2.9.6}$$

for all $k \in \mathbb{N}^d$ such that $k_1 + \cdots + k_d = n$, then $\beta_4(d) = \beta_1(d)$.

Proof. Set $r(\{O\}, \{O, x_1\}) = 1$. For $|A| \geq 1$, define $r(A, \cdot)$ recursively by equation (2.9.4). By induction, $|r(A, \cdot)| \leq 1$. By equation (2.9.3), $r(A, \cdot)$ makes up a collection of routing vectors. By equations (2.9.5) and (2.9.6), $f(A) = 1/c(|A|)$ for all $A \in \mathcal{C}_d$. Therefore, inequality (2.9.2) implies finite kinetic energy for $\beta > \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$. By Theorems 2.4.2 and 2.7.3, $\beta_4(d) \leq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}$. Combining this with Theorem 1.4.1b) and the fact that $\beta_1(d) \leq \beta_4(d)$ completes the proof. ■

Restrict attention to the case $d = 2$. Set $\rho(j) = c(j)/c(j+1)$. By the assumption that $\alpha_i(n; \cdot)$ is invariant under automorphisms of $\mathbb{B}^d$, it suffices to define $\alpha_1(n; k)$ for all $n \geq 1$ and for all $k \in \mathbb{N}^2$ such that $k_1 + k_2 = n - 1$. If, for all $n \geq 1$, $\alpha_i(n; \cdot)$ is a solution of

$$
\begin{aligned}
1 &= \alpha_1(n; (j, n-1-j)) + \alpha_2(n; (j, n-1-j)) \qquad |\alpha_i(n; (j, n-1-j))| \leq 1 \\
\rho(n) &= \alpha_1(n; (n-1, 0))\rho(n-1), \qquad (2.9.7) \\
\rho(n) &= \alpha_1(n; (j-1, n-j))\rho(j-1) + \alpha_2(n; (j, n-j-1))\rho(n-j-1)
\end{aligned}
$$

where $0 \leq j \leq n-1$, then Lemma 2.9.1 implies that $\beta_1(2) = \beta_4(2)$. By substituting $1 - \alpha_1(n; (j, n-1-j))$ for $\alpha_2(n; (j, n-j-1))$ in the final equation, solving equations (2.9.7) is equivalent to solving

$$
\begin{aligned}
\alpha_1(n; (j, n-1-j)) &\geq 0 \qquad \text{for } 0 \leq j \leq n-1, \\
\alpha_1(n; (n-1, 0)) &= \frac{\rho(n)}{\rho(n-1)} \qquad (2.9.8) \\
\alpha_1(n; (j-1, n-j)) &= \frac{\rho(n) - \rho(n-j-1) + \alpha_1(n; (j, n-1-j))\rho(n-1-j)}{\rho(j-1)},
\end{aligned}
$$

for $1 \leq j \leq n-1$ and $n \geq 1$.

Theorem 2.9.2 *The unique solution of equations (2.9.8) is*

$$\alpha_1(n;(j,n-1-j)) = \frac{(j+1)(2j+1)(3n-2j)}{n(n+1)(2n+1)}. \tag{2.9.9}$$

In particular, $\beta_1(2) = \beta_4(2)$.

Proof. Using the fact that $c(j) = \binom{2j}{j}/(2j+1)$, it follows that $\rho(j) = (j+2)/(4j+2)$ and therefore that

$$\frac{\rho(n)}{\rho(n-1)} = \frac{(2n-1)(n+2)}{(n+1)(2n+1)}.$$

Take $j = n-1$ in the righthand side of (2.9.9) to verify the base case. Assume that (2.9.9) holds for all m such that $j \le m \le n-1$. Then

$$\begin{aligned}
&\alpha_1(n;(j-1,n-j)) \\
&\quad= \frac{\rho(n)-\rho(n-j-1)+\alpha_1(n;(j,n-j-1))\rho(n-1-j)}{\rho(j-1)} \\
&\quad= \frac{4j-2}{j+1}\left(\frac{3(j+1)}{-8n^2+8jn+4j+2} + \frac{(j+1)(2j+1)(3n-2j)(n-j+1)}{n(n+1)(2n+1)(4n-4j-2)}\right) \\
&\quad= \frac{4j-2}{j+1}\frac{j(j+1)(3n-2(j-1))}{4n^3+6n+2n} \\
&\quad= \frac{(2j-1)j(3n-2(j-1))}{n(2n+1)(n+1)},
\end{aligned}$$

which proves the result. ∎

Theorem 2.9.2 together with Theorem 1.4.1a) imply that part d) of Theorem 1.4.1 holds. For $d \ge 3$, Lemma 2.9.1 reduces proving Conjecture 1.4.3 to proving that a solution to (2.9.3) and (2.9.6) exists. It is not hard to show that, disregarding the absolute bound of one requirement, there is a solution to (2.9.3) and

(2.9.6). The difficulty is that the solution is not unique. Therefore, verifying that a suitably bounded solution exists for all $n \in \mathbb{N}$ becomes more challenging. Chapter 3 is devoted to providing heuristic support for the existence of a solution that is absolutely bounded by one. We conclude this section with a proof of the existence of solutions that are not necessarily absolutely bounded by one.

Equations (2.9.3) and (2.9.6) make up a collection of linear algebra problems indexed by $\mathbb{N}$ that have the additional constraint that the solution is absolutely bounded by one. Farkas' Lemma provides an equivalent condition for proving that a system of linear equations has a solution. In the context of Linear Programming, Farkas' Lemma links dual programs. Essentially, the Duality Theorem is a translation of Farkas' Lemma into the language of Linear Programming. Almost any undergraduate text will discuss the connections between Farkas' Lemma and duality. See [12] Chapter 7 for example. Here the equivalent condition is verified for all $n \in \mathbb{N}$. Thus, disregarding the absolutely bound of one requirement, equations (2.9.3) and (2.9.6) have a solution for all $n \in \mathbb{N}$ and all $d \geq 2$.

Lemma 2.9.3 ***(Farkas).*** *Let A be an $m \times n$ matrix and b an m-vector. Then*

$$Ax = b \quad \textit{for some } x \in \mathbb{R}^n$$

if and only if

$$b^T y = 0 \quad \textit{for all} \quad y \in \mathbb{R}^m \quad \textit{such that} \quad A^T y = 0.$$

In our context, equations (2.9.3) and (2.9.6) determine the rows of the matrix A. To each equation, associate the (unordered) partition from which the equation originated. In particular,

$$\sum_{i=1}^{d} \alpha_i(n; k_1, \ldots, k_d) = 1 \longleftrightarrow (k_1, \ldots, k_d)$$

and

$$\sum_{i=1}^{d} \alpha_i(n; k_1, \ldots, k_i - 1, \ldots, k_d) \frac{c(k_i - 1)}{c(k_i)} = \frac{c(n)}{c(n+1)} \longleftrightarrow (k_1, \ldots, k_d).$$

It is not hard to see that each column of A has exactly two nonzero entries. This comes from the fact that each variable $\alpha_i(n; k_1, \ldots, k_d)$ appears once in the collection (2.9.3) and once in the collection (2.9.6). Thus, each row of A^T has exactly two nonzero entries. The entry that corresponds to the variable $\alpha_i(n; k_1, \ldots, k_d)$ in $A^T y$ is given by

$$\begin{aligned}(1 + |\{j : k_j = k_i + 1\}|) \frac{c(k_i)}{c(k_i + 1)} y(k_1, \ldots, k_i + 1, \ldots, k_d)& \\ + |\{j : k_j = k_i\}| y(k_1, \ldots, k_i, \ldots, k_d).&\end{aligned} \tag{2.9.10}$$

This easily implies that the dimension of the null space of A^T is at most one. To verify this, consider a graph in which the vertices correspond to the d-tuples in $\mathbb{N}^d$ with entries that add up to n or $n-1$. There is an edge between the vertices $(k_1, \ldots, k_i, \ldots, k_d)$ and $(l_1, \ldots, l_i, \ldots, l_d)$ if and only if $k_i = l_i$ for all except one index j and $|l_j - k_j| = 1$. In particular, the edges are in one-to-one correspondence with the equations in the collection (2.9.10). Since this graph is connected, there is at most one solution to equations (2.9.10). Let

$$y(k_1, \ldots, k_d) = (-1)^{k_1+\cdots+k_d} \frac{c(k_1)\cdots c(k_d)}{\prod_{j=1}^{d(k_1,\ldots,k_d)} |\{i : k_i = \delta_j\}|!}$$

where $d(k_1, \ldots, k_d)$ denotes the number of distinct entries in $(k_1, \ldots, k_d)$ and δ_j is the jth largest of these distinct entries. Then $y \in \mathrm{Null}(A^T)$ and since $y \neq 0$, y is a basis for $\mathrm{Null}(A^T)$. Using the fact that

$$\binom{d}{|\{i : k_i = \delta_1\}|, \ldots, |\{i : k_i = \delta_{d(k_1,\ldots,k_d)}\}|}$$

is the number of distinct orderings of the partition $(k_1, \ldots, k_d)$, it follows that

$$d!b^T y = (-1)^n \frac{c(n)}{c(n+1)} \sum_{k_1+\cdots+k_d=n} c(k_1)\cdots c(k_d) + (-1)^{n-1} \sum_{k_1+\cdots+k_d=n-1} c(k_1)\cdots c(k_d).$$

The Catalan recursion (recursion (2.5.2)) implies that

$$\frac{c(n)}{c(n+1)} \sum_{k_1+\cdots+k_d=n} c(k_1)\cdots c(k_d) - \sum_{k_1+\cdots+k_d=n-1} c(k_1)\cdots c(k_d) = 0.$$

Therefore, $d!b^T y = 0$. Farkas' Lemma implies that equations (2.9.3) and (2.9.6) have a solution. However, the lemma does not guarantee that there is a solution with the desired boundedness properties.

There is a form of Farkas' lemma that gives rise to the existence of a nonnegative solution. This involves verifying that $b^T y \geq 0$ for all vectors y such that $A^T y \geq 0$. Furthermore, the author believes that a nonnegative solution exists for $d \geq 3$. Nevertheless, attempts to use this version of Farkas' lemma to prove it have been unsuccessful, except in $d = 2$. Since Theorem 2.9.2 provides an explicit nonnegative solution when $d = 2$, using Farkas' Lemma to prove existence gives less information than Theorem 2.9.2. Therefore, Theorem 2.9.2 and its proof were presented here.

CHAPTER 3

The Continuous Problem

As previously noted, equations (2.9.3) and (2.9.6) make up a collection of linear algebra problems indexed by $\mathbb{N}$. Each problem has a distinct set of variables. Therefore, a solution to the $n = 5$ problem need not relate to a solution of the $n = 6$ problem. However, given the similarity of the equations it seems reasonable to expect that there exists a collection of solutions that are consistent in some sense. Any reasonable consistency condition will imply that the limit as n tends to infinity of $\alpha_i(n; \cdot)$ exists.

We investigate the limiting version of the equations (2.9.3) and (2.9.6). Under the limiting operation, equation (2.9.6) becomes a first order partial differential equation. It turns out that for all $d \geq 2$, the limiting version of (2.9.3) and (2.9.6) has a solution that is absolutely bounded by one. The existence of such a solution provides evidence that solutions to (2.9.3) and (2.9.6) exist that are absolutely bounded by one. Proving that such solutions exists, in turn proves Conjecture 1.4.3.

Here, a study of the limiting version of (2.9.3) and (2.9.6) is presented as support for Conjecture 1.4.3. In Section 3.1, the continuous problem is derived. In Section 3.2, the method used to find the solution is explained. The main idea is to assume that the solution can be expressed as a series and to devise a method for computing the coefficients. As one might expect, this approach becomes excessively complicated in general. However, the approach does provide an answer for small d and an educated guess for the general problem. An independent proof of Theorem 1.4.2 presented in Section 3.3.

3.1 The Derivation of the Continuous Problem

Assume that $\{\alpha(n, \cdot)\}_{n \in \mathbb{N}}$ is a set of solutions to the discrete problem such that

$$\alpha^*(x_1, \ldots, x_d) = \lim_{n \to \infty} \alpha_n^*(x_1, \ldots, x_d). \tag{3.1.1}$$

exists where $\alpha_n^*(x_1, \ldots, x_d) = \alpha_1(\lfloor nx_1 \rfloor + \cdots + \lfloor nx_d \rfloor + 1; \lfloor nx_1 \rfloor, \ldots, \lfloor nx_d \rfloor)$. By definition, $\alpha^*(x_1, \ldots, x_d)$ is symmetric in the variables $(x_2, \ldots, x_d)$. Furthermore, $\alpha^*(x_1, \ldots, x_d) = \alpha^*(ax_1, \ldots, ax_d)$ for all $a > 0$. Therefore,

$$\alpha^*(x_1, \ldots, x_d) = v\left(\frac{x_2}{x_1 + \cdots + x_d}, \ldots, \frac{x_d}{x_1 + \cdots + x_d}\right) \tag{3.1.2}$$

for some symmetric function v defined on the $d-1$ dimensional simplex $\mathbb{S}^{d-1}$. The limit of equation (2.9.3) is given by

$$\sum_{i=1}^{d} \alpha^*(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) = 1. \tag{3.1.3}$$

Letting $s_i = x_i/(x_1 + \cdots + x_d)$ for $1 \le i \le d$ and expressing equation (3.1.3) in terms of v,

$$v(s_2, \ldots, s_d) + \cdots + v(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d) + \cdots + v(s_1, \ldots, s_{d-1}) = 1. \tag{3.1.4}$$

If one simply takes the limit of equation (2.9.6), it collapses into equation (3.1.3). Therefore, first order information must be considered. By computing the first two coefficients of the power series centered at infinity,

$$\frac{c(n)}{c(n+1)} \sim \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}\left(1 + \frac{3}{2n}\right).$$

Expressing equation (2.9.6) in terms of α_n^* gives,

$$\begin{aligned}\sum_{i=1}^{d} \alpha_n^*(x_i - \tfrac{1}{n}, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)\frac{c(\lfloor n(x_i - \frac{1}{n})\rfloor)}{c(\lfloor nx_i \rfloor)}& \\ = \frac{c(\lfloor nx_1 \rfloor + \cdots + \lfloor nx_d \rfloor)}{c(\lfloor nx_1 \rfloor + \cdots + \lfloor nx_d \rfloor + 1)}&.\end{aligned} \tag{3.1.5}$$

Asymptotically, equation (3.1.5) is given by

$$\begin{aligned}\sum_{i=1}^{d} \alpha_n^*(x_i - \tfrac{1}{n}, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)\left(1 + \frac{3}{2n(x_i - \frac{1}{n})}\right)& \\ = 1 + \frac{3}{2n(x_1 + \cdots + x_d)} + o\left(\frac{1}{n}\right)&.\end{aligned}$$

As previously mentioned, first order information must be retained. Therefore, equation (3.1.6) will be multiplied by n. In order to prevent both sides from tending to infinity, (2.9.3) is subtracted from (3.1.6) before multiplication by n.

This gives

$$\sum_{i=1}^{d} n\left(\alpha_n^*(x_i - \tfrac{1}{n}, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) - \alpha_n^*(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)\right)$$
$$+ \frac{3}{2}\sum_{i=1}^{d} \frac{\alpha_n^*(x_i - \frac{1}{n}, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)}{x_i - \frac{1}{n}} = \frac{3}{2(x_1 + \cdots + x_d)} + o(1)\,.$$

Therefore, in the limit, equation (2.9.6) becomes

$$\sum_{i=1}^{d}\left(\frac{3}{2x_i} - \frac{\partial}{\partial x_i}\right)\alpha^*(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) = \frac{3}{2(x_1 + \cdots + x_d)}. \quad (3.1.6)$$

On the $d-1$ dimensional simplex, let

$$T_d v(s_1, \ldots, s_{d-1}) = \frac{v(s_1, \ldots, s_{d-1})}{1 - s_1 - \cdots - s_{d-1}} + \frac{2}{3}\sum_{i=1}^{d-1} s_i \frac{\partial}{\partial s_i} v(s_1, \ldots, s_{d-1}).$$

Multiplying equation (3.1.6) by $2(x_1 + \cdots + x_d)/3$ and expressing it in terms of $T_d v$,

$$\sum_{i=1}^{d} T_d v(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d) = 1. \quad (3.1.7)$$

For $w : \mathbb{S}^{d-1} \to \mathbb{R}$, let $L_d w : \partial\mathbb{S}^d \to \mathbb{R}$ be defined by

$$L_d w(s_1, \ldots, s_d) = \sum_{i=1}^{d} w(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d). \quad (3.1.8)$$

Equations (3.1.4) and (3.1.7) can be expressed in terms of L_d as

$$L_d v(s_1, \ldots, s_d) = 1 \qquad \text{and} \qquad L_d T_d v(s_1, \ldots, s_d) = 1 \quad (3.1.9)$$

respectively. The next proposition which summarizes the statement of the continuous problem has been proved.

Proposition 3.1.1 *If $\alpha^* : \mathbb{R}_+^d \to \mathbb{R}$ is symmetric in the variables $x_2, \ldots, x_d$ and satisfies equations (3.1.3) and (3.1.6), then $v : \mathbb{S}^{d-1} \to \mathbb{R}$ defined by equation (3.1.2) is a symmetric solution of equations (3.1.9). Conversely, if $v : \mathbb{S}^{d-1} \to \mathbb{R}$ is symmetric and satisfies equations (3.1.9), then $\alpha^* : \mathbb{R}_+^d \to \mathbb{R}$ defined by equation (3.1.2) is symmetric in the variables $x_2, \ldots, x_d$ and satisfies equations (3.1.3) and (3.1.6).*

3.2 The Method for Finding a Solution

The method used to actually find the solution is presented in this section. The strategy is to express a candidate solution as a series with unknown coefficients and to use the partial differential equation to determine the coefficients. The approach is demonstrated in $d = 3$ and only the main ideas are presented. In Section 3.3, a complete proof of Theorem 1.4.2 is given that is independent of the approach taken here.

The goal is to find $v(s,t)$ such that

$$v(s,t) + v(1-s-t,t) + v(1-s-t,s) = 1$$

$$T_3 v(s,t) + T_3 v(1-s-t,t) + T_3 v(1-s-t,s) = 1.$$

For any such $v(s,t)$, $v(s,t) = u(s,t) + 1/3$ for some $u(s,t)$ that satisfies

$$u(s,t) + u(1-s-t,t) + u(1-s-t,s) = 0. \tag{3.2.1}$$

A collection of symmetric polynomials that satisfy equation (3.2.1) is given by

$$\begin{aligned} u_{n,m}(s,t) &= (1-s-t)^n(st)^n\left((1-t-2s)t^m + (1-s-2t)s^m\right) \\ &= (1-s-t)^{n+1}(st)^n(s^m+t^m) - (1-s-t)^n(st)^{n+1}(s^{m-1}+t^{m-1}) \end{aligned}$$

where $m, n \in \mathbb{N}$. Consider $u(s,t) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \omega_{n,m} u_{n,m}(s,t)$ where $\omega_{n,m} \in \mathbb{R}$. The goal is to choose $\omega_{n,m}$ such that $T_3(u+1/3)(s,t) - 1/3$ also satisfies equation (3.2.1).

Since $T_3 u_{n,m}(s,t)$ is not expressible in terms of the collection $\{u_{n,m}(s,t)\}_{n,m\in\mathbb{N}}$, some symmetric polynomials are added to the collection. Let

$$p_{n,m}(s,t) = (1-s-t)^n(st)^n(s^m+t^m).$$

The collection $\{u_{n,m}(s,t), p_{n,m}(s,t)\}_{n,m\in\mathbb{N}}$ spans the set of all symmetric polynomials in two variables. It turns out that

$$\begin{aligned} T_3 u_{n,m}(s,t) =& \frac{2(3n+m+1)}{3} u_{n,m}(s,t) \\ &- \left(\frac{3-2n}{6}\right)\left(u_{n-1,m+2}(s,t) - u_{n-1,m+1}(s,t)\right) \\ &+ \left(\frac{4-4n}{3}\right) p_{n,m}(s,t) \\ &- \left(\frac{3-2n}{6}\right)\left(p_{n-1,m+1}(s,t) - 2p_{n-1,m+2}(s,t) + p_{n-1,m+3}(s,t)\right). \end{aligned}$$

Also,

$$T_3(1/3)(s,t) - 1/3 = \frac{s+t}{3(1-s-t)} = \frac{p_{-1,2}(s,t) - p_{-1,3}(s,t) - u_{-1,2}(s,t)}{6}.$$

Recall that the objective is to choose $\omega_{n,m}$ such that $T_3(u+1/3)(s,t)-1/3$ satisfies equation (3.2.1). In other words, the coefficient of $p_{n,m}(s,t)$ in $T_3(u+1/3)(s,t)-1/3$ should be zero for all m and n. If $\kappa_{m,n}$ denotes the coefficient of $p_{n,m}(s,t)$ in $T_3(u+1/3)-1/3$, it follows that

$$\kappa(n,m)=\begin{cases}\frac{1}{6}+\frac{8\omega(-1,2)}{3}-\frac{\omega(0,1)}{2}+\omega(0,0)-\frac{\omega(0,-1)}{2} & \text{if } (n,m)=(-1,2),\\ -\frac{1}{6}+\frac{8\omega(-1,3)}{3}-\frac{\omega(0,2)}{2}+\omega(0,1)-\frac{\omega(0,0)}{2} & \text{if } (n,m)=(-1,3).\end{cases}$$

Otherwise,

$$\kappa(n,m)=\left(\frac{4-4n}{3}\right)\omega(n,m) + \left(\frac{2n-1}{3}\right)\left(\frac{\omega(n+1,m-1)}{2}-\omega(n+1,m-2)+\frac{\omega(n+1,m-3)}{2}\right).$$

Setting $\kappa(n,m)=0$ implies that

$$\omega(n,m)=\begin{cases}\frac{1}{3} & \text{if } n=0 \text{ and } m\geq 1,\\ \frac{4(m+1)(m+2)}{3} & \text{if } n=1 \text{ and } m\geq 0,\\ 0 & \text{otherwise.}\end{cases}$$

By summing the series that defines $u(s,t)$,

$$u(s,t)=\frac{(1-t-2s)t}{3(1-t)}+\frac{(1-s-2t)s}{3(1-s)} + \frac{8(1-t-2s)st(1-s-t)}{3(1-t)^3}+\frac{8(1-s-2t)st(1-s-t)}{3(1-s)^3}. \tag{3.2.2}$$

If one generalizes this approach and repeats the procedure for $d=4$, the solution is given by

$$
\begin{aligned}
&u(r,s,t) = \\
&\frac{(1-r-s-2t)(r+s)}{8(1-r-s)} + \frac{(1-r-t-2s)(r+t)}{8(1-r-t)} + \frac{(1-s-t-2r)(s+t)}{8(1-s-t)} \\
&+ \frac{(1-r-s-t)(1-r-s-2t)(r+s)t}{(1-r-s)^3} + \frac{(1-r-s-t)(1-r-t-2s)(r+t)s}{(1-r-t)^3} \\
&+ \frac{(1-r-s-t)(1-s-t-2r)(s+t)r}{(1-s-t)^3}.
\end{aligned} \tag{3.2.3}
$$

Comparing the $d = 3$ and $d = 4$ solutions suggests a pattern. Since computing the coefficients is complicated in general, it is more convenient to verify that the candidate solution satisfies equations (3.1.9).

3.3 A Solution to the Continuous Problem

In this section, the pattern suggested by equations (3.2.2) and (3.2.3) is shown to satisfy equations (3.1.9). The proof itself heavily exploits the structure of the solution and thus reveals the properties of the solution that enable it to satisfy equations (3.1.9).

Definition 3.3.1 *For $u : \mathbb{S}^{d-1} \to \mathbb{R}$, u is homogeneous with respect to L_d if $L_d u = 0$. Denote the set of all symmetric functions that are homogeneous with respect to L_d by $\mathcal{H}_d$.*

The next proposition is an immediate consequence of this definition.

Proposition 3.3.2 *If $u \in \mathcal{H}_d$ and $T_d(u + 1/d) - 1/d \in \mathcal{H}_d$, then $v = u + 1/d$ is a symmetric solution of equations (3.1.9).*

Let φ be the projection of $\mathbb{S}^{d-1}$ onto $\mathbb{S}^2$ defined by

$$\varphi(s_1, s_2, \ldots, s_{d-1}) = (s_1, s_2 + \cdots + s_{d-1}). \tag{3.3.1}$$

Given a function $f : \mathbb{S}^2 \to \mathbb{R}$, let $S_d f$ be the symmetrized extension of f to $\mathbb{S}^{d-1}$ defined by

$$S_d f(s_1, \ldots, s_{d-1}) = \sum_{i=1}^{d-1} f \circ \varphi(s_i, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{d-1}).$$

It is immediate that S_d is a linear operator and that $S_d f$ is symmetric. The class of functions that will be considered here are all symmetrized extensions. In particular, we consider $u \in \mathcal{H}_d$ such that $u = S_d f$ some $f : \mathbb{S}^2 \to \mathbb{R}$. By restricting attention to this class, we can view our solution as a sum of functions of two variables. There is a simple criterion for functions $f : \mathbb{S}^2 \to \mathbb{R}$ that implies that $S_d f \in \mathcal{H}_d$.

Definition 3.3.3 *Given $f : \mathbb{S}^2 \to \mathbb{R}$, we say that f is cancelative if $f(s,t) + f(1 - s - t, t) = 0$ for all $(s,t) \in \mathbb{S}^2$.*

Proposition 3.3.4 *If $f : \mathbb{S}^2 \to \mathbb{R}$ is cancelative, then $S_d f \in \mathcal{H}_d$.*

Proof. By definition, $s_1 + \cdots + s_d = 1$. Thus, for $1 \le i < j \le d$,

$$\begin{aligned}
&\varphi_1(s_i, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d) \\
&\qquad = 1 - \varphi_1(s_j, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d) \\
&\qquad\quad - \varphi_2(s_j, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d)
\end{aligned} \tag{3.3.2}$$

$$\begin{aligned}
&\varphi_2(s_i, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d) \\
&\qquad = \varphi_2(s_j, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d).
\end{aligned} \tag{3.3.3}$$

where φ_i denotes the ith coordinate of φ. Combining equations (3.3.2) and (3.3.3) with the fact that f is cancelative implies that

$$\begin{aligned}
&f \circ \varphi(s_i, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d) \\
&\qquad + f \circ \varphi(s_j, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{j-1}, s_{j+1}, \ldots, s_d) = 0.
\end{aligned} \tag{3.3.4}$$

By definition,

$$\begin{aligned}
L_d S_d f &= \sum_{i=1}^{d} S_d f(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d) \\
&= \sum_{i=2}^{d} \sum_{j<i} f \circ \varphi(s_j, \ldots, s_{j-1}, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d) \\
&\quad + \sum_{i=1}^{d-1} \sum_{i<j} f \circ \varphi(s_j, \ldots, s_{i-1}, s_{i+1}, \ldots, s_1, s_{j+1}, \ldots, s_d).
\end{aligned}$$

By switching the order of the second pair of summations, equation (3.3.4) implies that

$$\begin{aligned}
L_d S_d f &= \sum_{i=2}^{d} \sum_{j<i} f \circ \varphi(s_j, \ldots, s_{j-1}, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d) \\
&\quad + \sum_{i=2}^{d} \sum_{j<i} f \circ \varphi(s_i, \ldots, s_{j-1}, s_{j+1}, \ldots, s_1, s_{i+1}, \ldots, s_d) \\
&= \sum_{i=2}^{d} \sum_{j<i} (f \circ \varphi(s_j, \ldots, s_{j-1}, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_d) \\
&\qquad + f \circ \varphi(s_i, \ldots, s_{j-1}, s_{j+1}, \ldots, s_1, s_{i+1}, \ldots, s_d)) \\
&= 0,
\end{aligned}$$

completing the proof. ∎

Two examples of cancelative functions are

$$(1 - t - 2s) \qquad \text{and} \qquad s(1 - s - t)(1 - t - 2s).$$

These two examples will be the main building blocks for the solution to equations (3.1.9). Note that if either example is multiplied by a function that depends only on the variable t, then the resulting function is also cancelative. In particular, if

$$f(s,t) = \frac{(1 - t - 2s)t}{(1 - t)} \qquad \text{and} \qquad g(s,t) = \frac{(1 - t - 2s)s(1 - s - t)t}{(1 - t)^3}, \tag{3.3.5}$$

then $S_d f$ and $S_d g$ are elements of $\mathcal{H}_d$. Furthermore, $S_d(a_d f + b_d g)$ is an element of $\mathcal{H}_d$ for any real constants a_d and b_d. Our goal is to choose a_d and b_d such that $T_d(a_d f + b_d g + 1/d) - 1/d \in \mathcal{H}_d$.

Proposition 3.3.5 *For all* $f : \mathbb{S}^2 \to \mathbb{R}$, $T_d(f \circ \varphi(s_1, \ldots, s_{d-1})) = (T_3 f) \circ \varphi(s_1, \ldots, s_{d-1})$. *In particular,* $T_d S_d f = S_d T_3 f$.

Proof. By the chain rule,

$$s_1 \frac{\partial}{\partial s_1} f \circ \varphi(s_1, \ldots, s_{d-1}) = s_1 \frac{\partial f}{\partial s} (\varphi(s_1, \ldots, s_{d-1}))$$
$$s_i \frac{\partial}{\partial s_i} f \circ \varphi(s_1, \ldots, s_{d-1}) = s_i \frac{\partial f}{\partial t} (\varphi(s_1, \ldots, s_{d-1})) \qquad i \neq 1.$$

Therefore,

$$\sum_{i=1}^{d-1} s_i \frac{\partial}{\partial s_i} f \circ \varphi(s_1, \ldots, s_{d-1})$$
$$= \varphi_1(s_1, \ldots, s_{d-1}) \frac{\partial f}{\partial s} (\varphi(s_1, \ldots, s_{d-1})) + \varphi_2(s_1, \ldots, s_{d-1}) \frac{\partial f}{\partial t} (\varphi(s_1, \ldots, s_{d-1})),$$

completing the proof. ∎

As a consequence of Proposition 3.3.5 and linearity of both S_d and T_d,

$$T_d S_d(a_d f + b_d g) = a_d S_d T_3 f + b_d S_d T_3 g.$$

Therefore, it is enough to compute $T_3 f$ and $T_3 g$. In light of Proposition 3.3.4, the next objective is to collect all cancelative parts of $T_3 f$ and $T_3 g$.

Proposition 3.3.6 *If* f *and* g *are defined by (3.3.5), then*

$$T_3 f(s,t) = \frac{4t}{3(1-t)} - \frac{t}{1-s-t} + \frac{2}{3}\left(2 + \frac{t}{1-t}\right) f(s,t) \qquad (3.3.6)$$

$$T_3 g(s,t) = \frac{-st}{3(1-t)^2} + \frac{2}{3}\left(4 + \frac{3t}{1-t}\right) g(s,t). \qquad (3.3.7)$$

Proof. We have

$$s\frac{\partial f}{\partial s}(s,t) = \frac{-2st}{1-t} \tag{3.3.8}$$

$$t\frac{\partial f}{\partial t}(s,t) = \frac{t(1-t-2s)}{1-t} + \frac{-t^2}{1-t} + \frac{t^2(1-t-2s)}{(1-t)^2} \tag{3.3.9}$$

The first term in equation (3.3.9) is $f(s,t)$. By combining the second term in equation (3.3.9) with the right hand side of equation (3.3.8) and adding and subtracting $t/(1-t)$, another copy of $f(s,t)$ can be obtained. The final term in equation (3.3.9) is simply $f(s,t)$ scaled by a function that depends only on the variable t. Thus,

$$s\frac{\partial f}{\partial s}(s,t) + t\frac{\partial f}{\partial t}(s,t) = \left(2 + \frac{t}{1-t}\right) f(s,t) - \frac{t}{1-t}.$$

Observing that

$$\frac{f(s,t)}{1-s-t} = \frac{2t}{1-t} - \frac{t}{(1-s-t)}$$

gives

$$\begin{aligned} T_3 f(s,t) &= \frac{2t}{1-t} - \frac{t}{(1-s-t)} + \frac{2}{3}\left(\left(2 + \frac{t}{1-t}\right) f(s,t) - \frac{t}{1-t}\right) \\ &= \frac{4t}{3(1-t)} - \frac{t}{(1-s-t)} + \frac{2}{3}\left(2 + \frac{t}{1-t}\right) f(s,t), \end{aligned}$$

establishing equation (3.3.6).

For $g(s,t)$,

$$s\frac{\partial g}{\partial s}(s,t) = \frac{s(1-t-2s)(1-s-t)t}{(1-t)^3} + \frac{-2s^2(1-s-t)t}{(1-t)^3} + \frac{-s^2(1-t-2s)t}{(1-t)^3} \tag{3.3.10}$$

$$\begin{aligned} t\frac{\partial g}{\partial t}(s,t) = &\frac{-ts(1-s-t)t}{(1-t)^3} + \frac{-t(1-t-2s)st}{(1-t)^3} + \frac{t(1-t-2s)s(1-s-t)}{(1-t)^3} \\ &+ \frac{3t(1-t-2s)s(1-s-t)t}{(1-t)^4} \end{aligned} \tag{3.3.11}$$

In a similar manner as with $f(s,t)$, combine the second and third terms in equation (3.3.10) with the first and second terms in (3.3.11) respectively to obtain

$$s\frac{\partial g}{\partial s}(s,t) + t\frac{\partial g}{\partial t}(s,t) = \left(4 + \frac{3t}{1-t}\right) g(s,t) - \frac{s(2-2t-3s)t}{(1-t)^3}$$

Since

$$\frac{g(s,t)}{1-s-t} - \frac{2s(2-2t-3s)t}{3(1-t)^3} = \frac{-st}{3(1-t)^2},$$

equation (3.3.7) holds. ∎

Proposition 3.3.6 decomposes $T_3 f$ and $T_3 g$ into cancelative and noncancelative components. Denote the noncancelative terms by

$$\varepsilon_1(s,t) = \frac{4t}{3(1-t)}, \qquad \varepsilon_2(s,t) = \frac{-t}{1-s-t}, \qquad \text{and} \qquad \varepsilon_3(s,t) = \frac{-st}{3(1-t)^2}.$$

Recall our ultimate goal, to choose a_d and b_d such that $T_d S_d(a_d f + b_d g) + T_d(1/d) - 1/d$ is an element of $\mathcal{H}_d$. Since

$$T_d(1/d) - 1/d = \frac{s_1 + \cdots + s_{d-1}}{d(1 - s_1 - \cdots - s_{d-1})}$$

and

$$S_d \varepsilon_2 = \frac{-(d-2)(s_1 + \cdots + s_{d-1})}{1 - s_1 - \cdots - s_{d-1}},$$

it is natural to choose a_d such that $S_d a_d \varepsilon_2$ cancels $T_d(1/d) - 1/d$. In particular, $a_d = 1/((d-2)d)$ so that

$$T_d(1/d) - 1/d + S_d a_d \varepsilon_3 = 0. \tag{3.3.12}$$

With only ε_1 and ε_3 remaining, b_d is chosen such that $a_d\varepsilon_1 + b_d\varepsilon_3$ is cancelative. Setting $b_d = 8a_d$ gives

$$a_d\varepsilon_1(s,t) + 8a_d\varepsilon_3(s,t) = a_d\frac{4t(1-t-2s)}{3(1-t)^2} = a_d\frac{4}{3(1-t)}f(s,t) \qquad (3.3.13)$$

which is cancelative.

Theorem 3.3.7 *Let $h(s,t) = f(s,t) + 8g(s,t)$ where f and g are defined by equations (3.3.5). Then $S_d a_d h + 1/d$ is a symmetric solution to equations (3.1.9).*

Proof. Since $h(s,t)$ is cancelative, Proposition 3.3.4 implies that $S_d a_d h \in \mathcal{H}_d$. By Proposition 3.3.6 and equation (3.3.13),

$$T_3 a_d h(s,t) = a_d\frac{2(4-t)}{3(1-t)}h(s,t) + a_d\varepsilon_2(s,t). \qquad (3.3.14)$$

Since $h(s,t)$ is cancelative, Proposition 3.3.4, Proposition 3.3.5, and equation (3.3.12) imply that $T_d(S_d a_d h + 1/d) - 1/d$ is an element of $\mathcal{H}_d$. By Proposition 3.3.2, the assertion holds. ∎

Theorem 3.3.8 *$S_d a_d h + 1/d$ is absolutely bounded by one.*

Proof. We have

$$\frac{\partial h}{\partial s}(s,t) = \frac{6t((1-t)^2 - 8(1-t)s + 8s^2)}{(1-t)^3}.$$

Therefore, the maximum and minimum occur at

$$s_{\max} = \frac{(1-t)(2-\sqrt{2})}{4} \qquad \text{and} \qquad s_{\min} = \frac{(1-t)(2+\sqrt{2})}{4}$$

respectively. Since

$$h\left(\frac{(1-t)(2-2\sqrt{2})}{4}, t\right) = \sqrt{2}t \qquad \text{and} \qquad h\left(\frac{(1-t)(2+2\sqrt{2})}{4}, t\right) = -\sqrt{2}t,$$

It follows that

$$|h \circ \varphi(s_1, \ldots, s_{d-1})| \leq \sqrt{2} \qquad \text{on} \qquad \mathbb{S}^{d-1}.$$

Since $S_d h$ has $d-1$ terms of the form $h \circ \varphi$,

$$|S_d a_d h + \frac{1}{d}| \leq \frac{(d-1)\sqrt{2}}{(d-2)d} + \frac{1}{d}$$

which is bounded by one provided $d \geq 4$. Since $S_3 h(s,t) = h(s,t) + h(t,s)$, it is possible to use the better bound of

$$|h(s,t) + h(t,s)| \leq \sqrt{2}(s+t) \leq \sqrt{2}.$$

Thus,

$$|S_3 a_3 h + \frac{1}{3}| \leq \frac{\sqrt{2}+1}{3},$$

as desired. ∎

Theorem 1.4.2 follows from Theorem 3.3.7, Theorem 3.3.8, and Proposition 3.1.1. Presumably, a suitably bounded solution to the discrete problem exists that has a structure analogous to the structure of the solution to the partial differential equation. Our attempts to exploit this structure have failed. Nevertheless, we believe that (2.9.3) and (2.9.6) has a solution for all $n \in \mathbb{N}$.

CHAPTER 4

Critical Exponents

The focus of this chapter is the behavior of the survival probability, the expected extinction time, and the susceptibility as functions of β. Of particular interest will be the behavior of these functions near the critical values. In Section 4.3, the continuity properties of the survival probability are examined. Next we turn our attention to the problem of bounding the survival probability from above and below. The bounds derived in Sections 4.4 and 4.5 lead to a proof of Theorem 1.4.4. In the final section of this chapter, explicit formulas are derived for the expected extinction time and the susceptibility.

4.1 The Dirichlet and Thompson's Principles

The Dirichlet principle and Thompson's principle provide powerful tools for describing the behavior of the survival probability. These principles apply in the setting of a reversible Markov chain. The Dirichlet principle states that the probability that the Markov chain escapes from some fixed subset of the state space before returning to the initial state is expressible as an infimum of a certain variational functional over all functions in some class. Likewise, Thompson's principle expresses this same probability as a supremum of an energy functional over all functions in some class. Furthermore, there is a unique function that optimizes these functionals. The precise statements of these principles are as follows.

Let X_t be a reversible Markov chain with state space S, stationary measure π, and transition rates $q(x,y)$. For any subset R of the state space S, let

$$\tau_R = \inf\{t : X_t \in R\} \qquad \text{and} \qquad \tau_R^+ = \inf\{t > \tau_{R^c} : X_t \in R\}.$$

Given a function $h : S \to [0,1]$, let $\Phi(h)$ be the *Dirichlet form* evaluated at h:

$$\Phi(h) = \frac{1}{2}\sum_{x,y} \pi(x)q(x,y)(h(y) - h(x))^2.$$

Given a subset R of the state space S and $x \in S \setminus R$, let

$$\mathcal{H}_x^R = \{h : S \to [0,1] : h(x) = 0, h(y) = 1 \text{ for all } y \in R\}.$$

Theorem 4.1.1 ***(The Dirichlet Principle)***. *Provided $P^x(\tau_R < \infty) = 1$,*

$$\pi(x)q(x)P^x(\tau_R \le \tau_x^+) = \inf_{h \in \mathcal{H}_x^R} \Phi(h),$$

where $q(x) = \sum_{y \neq x} q(x,y)$. Furthermore,

$$h(y) = P^y(\tau_R \le \tau_x)$$

is the unique function that is harmonic on $S \setminus (R \cup x)$ with the stated boundary conditions and the infimum is attained by this function.

Recall that a function is said to be harmonic on some subset of the state space U if

$$h(x) = \sum_{y \neq x} \frac{q(x,y)}{q(x)} h(y) \qquad \text{for all } x \in U.$$

In other words, h satisfies an averaging property on U. For an irreducible subset U of the state space, this averaging property implies that if the function h attains its maximum or minimum value in U, then h is constant on U. Consequently, specifying the values of h on U complement and requiring harmonicity on U determines h, provided the Markov chain hits U complement with probability one. It turns out that $\Phi(h)$ is minimal on $\mathcal{H}_x^R$ if and only if h is harmonic on $S \setminus (R \cup x)$. A proof of the Dirichlet principle can be found in Liggett [17] .

The Dirichlet principle can be stated in a dual form known as Thompson's principle. Given an anti-symmetric function $w : S \times S \to \mathbb{R}$, let $\mathcal{K}(w)$ denote the *kinetic energy* of w:

$$\mathcal{K}(w) = \frac{1}{2} \sum_{x,y} \frac{w^2(x,y)}{\pi(x) q(x,y)}.$$

Given a subset R of the state space S and $x \in S \setminus R$, let

$$\begin{aligned} \mathcal{W}_x^R = \{ & w : S \times S \to \mathbb{R} : w(y,z) = -w(z,y), \ y, z \in S; \\ & \sum_y w(x,y) = 1; \text{ and } \sum_y w(z,y) = 0, z \notin R \cup x \}. \end{aligned}$$

Such a function w is said to be a *unit flow* from x to R.

Theorem 4.1.2 ***(Thompson's Principle).*** *Provided* $P^x(\tau_R < \infty) = 1$,

$$\sup_{w \in \mathcal{W}_x^R} \frac{1}{\mathcal{K}(w)} = \pi(x)q(x)P^x(\tau_R \leq \tau_x^+),$$

where $q(x) = \sum_{y \neq x} q(x,y)$. *Furthermore, the unit flow given by*

$$\begin{aligned} w(y,z) =& \mathbb{E}^x \left(\textit{number of one step transitions from } y \textit{ to } z \textit{ before time } \tau_R \right) \\ &- \mathbb{E}^x \left(\textit{number of one step transitions from } z \textit{ to } y \textit{ before time } \tau_R \right). \end{aligned}$$

attains the supremum.

Not surprisingly, the optimal unit flow is related to the harmonic function that appears in the Dirichlet principle. To see this, define a path from $y \notin R$ to R to be a sequence $\{y_i\}_{i=0}^m$ of states in the Markov chain such that $y_0 = y$, $q(y_i, y_{i+1}) > 0$, and $\{y_i\}_{i=0}^m \cap R = y_m$. The optimal flow satisfies

$$\sum_{i=0}^{n-1} \frac{w(x_i, x_{i+1})}{\pi(x_i)q(x_i, x_{i+1})} = \sum_{i=0}^{m-1} \frac{w(y_i, y_{i+1})}{\pi(y_i)q(y_i, y_{i+1})},$$

for all pairs of paths to R such that $x_0 = y_0$. Therefore, if $x_0 = x$ and $x_k = y$, then the function

$$h(y) = \sum_{i=0}^{k-1} \frac{w(x_i, x_{i+1})}{\pi(x_i)q(x_i, x_{i+1})}$$

is well defined. Furthermore, the incompressibility property of $w(y,z)$ implies that h is harmonic. After normalizing h so that it takes the value one on R, we see that the optimal flow is related to the harmonic function from the Dirichlet principle by the equation

$$w(y,z) = \frac{\pi(y)q(y,z)(h(z) - h(y))}{\pi(x)q(x)P^x(\tau_R \leq \tau_x^+)}$$

The book by Doyle and Snell [8] is the standard reference for this topic.

4.2 The Shape Chain

In order to apply the Dirichlet principle and Thompson's principle to the uniform model, a related reversible Markov chain is introduced called the *shape chain*. As motivation for the definition of this Markov chain, observe that the issue of whether or not the uniform model avoids absorption into the empty set is independent of the location of the of the occupied set. Furthermore, the evolution of the uniform model depends only on the 'shape' of the occupied set. So, it seems reasonable to identify isomorphic occupied sets and record the shape rather than the location of the occupied set. This allows a transition from the empty set to the singleton to be introduced while preserving reversibility.

More formally, the shape chain is defined as follows. An automorphism of a graph $G = (V, E)$ is a bijection $\phi : V \to V$ such that there is an edge $e_1 \in E$ between the vertices x and y if and only if there is an edge $e_2 \in E$ between the vertices $\phi(x)$ and $\phi(y)$. Let $\mathrm{Aut}(\mathbb{T}^d)$ be the set of all automorphisms of $\mathbb{T}^d$. Configurations A and B are said to be equivalent if there exists $\phi \in \mathrm{Aut}(\mathbb{T}^d)$ such that $\phi(B) = A$. We write $A \sim B$ to indicate that A and B are equivalent. The relation $\sim$ defines an equivalence relation on the set of all configurations. Let

$\hat{A} = \{B : B \sim A\}$ and

$$\hat{\mathcal{S}} = \hat{\emptyset} \cup \{\hat{A} : A \text{ is a finite connected subset of } \mathbb{T}^d\}.$$

Roughly speaking, $\hat{\mathcal{S}}$ denotes the set of all finite connected shapes that can be embedded into $\mathbb{T}^d$. It will be convenient to consider the Markov chain $\hat{A}_t$ induced on $\hat{\mathcal{S}}$ by the dynamics of the uniform model. In particular, for $\hat{A} \neq \hat{\emptyset}$

$$\hat{q}(\hat{A}, \hat{B}) = \sum_{\{x: A_x \sim B\}} c(x, A)$$

where $A \in \hat{A}$ is fixed and A_x is $A \cup x$ if $x \notin A$ and $A \setminus x$ if $x \in A$. Since $\hat{q}(\hat{A}, \cdot)$ depends on A only through its equivalence class, the transition rates are well defined. We refer to $\hat{A}_t$ as the shape chain. In order to make the shape chain irreducible, a transition from $\hat{\emptyset}$ to the singleton $\hat{O}$ is introduced at rate β.

The shape chain is reversible with respect to the measure

$$\hat{\pi}(\hat{A}) = \frac{M(\hat{A})\beta^{|\hat{A}|}}{|\hat{A}|},$$

where $M(\hat{A}) = |\{A \in \hat{A} : O \in A\}|$ and $|\hat{A}|$ is the number of vertices in $A \in \hat{A}$. In order to prove this, it suffices to show that the detailed balance equations hold:

$$\hat{\pi}(\hat{A})\hat{q}(\hat{A}, \hat{B}) = \hat{\pi}(\hat{B})\hat{q}(\hat{B}, \hat{A}) \tag{4.2.1}$$

for all $\hat{A}, \hat{B} \in \hat{\mathcal{S}}$. Without loss of generality, $|\hat{B}| \geq |\hat{A}|$. If either the left or the right hand side of (4.2.1) is nonzero, then there exist $A \in \hat{A}$ and $B \in \hat{B}$ such that $A \cup x = B$ for some $x \in \mathbb{T}^d$. Thus, proving that equation (4.2.1) holds is equivalent to proving that

$$\frac{M(\hat{A})|\{D \sim B : D \supset A\}|}{|\hat{A}|} = \frac{M(\hat{B})|\{C \sim A : C \subset B\}|}{|\hat{B}|} \tag{4.2.2}$$

for all finite, connected subsets A and B containing O such that $A \cup x = B$. Liggett proves that equation (4.2.2) holds for all finite (not necessarily connected) subsets: see equation (3.8) in [18].

By definition, the shape chain starting from the singleton $\hat{A}_t^{\hat{O}}$ and the uniform model starting from the origin η_t^O can be coupled such that $\eta_t^O \in \hat{A}_t^{\hat{O}}$ for all times $t \leq \tau_{\hat{\emptyset}}$. Thus, the problem of determining the asymptotic behavior of $P(\eta_t^O \neq \emptyset \; \forall \; t)$ as β decreases to β_2 is equivalent to determining the asymptotic behavior of $P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty)$. Also, note that $P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty)$ can be expressed as a limit of escape probabilities. To see this, fix a sequence $\{\hat{\mathcal{S}}_N\}$ of subsets of $\hat{\mathcal{S}}$ that has the properties that $\hat{\emptyset} \in \hat{\mathcal{S}} \setminus \hat{\mathcal{S}}_N$ increases to $\hat{\mathcal{S}}$, $\hat{\emptyset} \notin \hat{\mathcal{S}}_N$, and $P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} < \infty) = 1$ for each $N \in \mathbb{N}$. Since $P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} \leq \tau_{\hat{\emptyset}}^+) = P^{\hat{O}}(\tau_{\hat{\mathcal{S}}_N} \leq \tau_{\hat{\emptyset}})$,

$$P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty) = \lim_{N \to \infty} P^{\hat{O}}(\tau_{\hat{\mathcal{S}}_N} \leq \tau_{\hat{\emptyset}}) = \lim_{N \to \infty} P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} \leq \tau_{\hat{\emptyset}}^+).$$

Therefore,

$$P(\eta_t^O \neq \emptyset \; \forall \; t) = \lim_{N \to \infty} P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} \leq \tau_{\hat{\emptyset}}^+)$$

Since the probabilities $P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} \leq \tau_{\hat{\emptyset}}^+)$ can be expressed in terms of the Dirichlet principle and Thompson's principle, this framework provides a strategy for estimating the survival probability.

4.3 Continuity of the Survival Probability

Before proceeding to estimate the survival probability, we discuss the continuity properties of the function

$$s(\beta) = P^O(\eta_t \neq \emptyset \ \forall \ t).$$

The fact that this function is right continuous is rather easy to establish as it is upper semi-continuous and increasing. Furthermore, it is continuous on $[0, \beta_2)$ as a matter of definition. Therefore, the main issue is to establish left continuity above β_2. Generally speaking, continuity from the left at β_2 is the most difficult to prove. Since we expect that $\beta_1 = \beta_4$ for all d and $s(\beta_1) = 0$, continuity from the left at β_2 would follow from a proof of Conjecture 1.4.3. Since it is known that $\beta_1 = \beta_4$ on the binary tree, it will follow from the arguments given here that $s(\beta)$ is continuous on the binary tree.

Here the continuity of $s(\beta)$ for β sufficiently large is established. We begin by reviewing the proof of right continuity. Then continuity for β sufficiently large is obtained as an application of the Dirichlet principle.

Definition 4.3.1 *A function $f : \mathbb{R} \to \mathbb{R}$ is said to be upper semi-continuous if $\{x : f(x) < a\}$ is open for all $a \in \mathbb{R}$.*

Proposition 4.3.2

i) If $f : \mathbb{R} \to \mathbb{R}$ is upper semi-continuous, then $\limsup_{x \to y} f(x) \le f(y)$.

ii) If $f_i : \mathbb{R} \to \mathbb{R}$ is upper semi-continuous for each i in some index set $\mathbb{I}$, then $\inf_i f_i(x) = f(x)$ *is upper semi-continuous.*

iii) If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then $f : \mathbb{R} \to \mathbb{R}$ is upper semi-continuous.

Properties i), ii), and iii) in Proposition 4.3.2 are immediate consequences of the definition of upper semi-continuity.

Lemma 4.3.3 *The function $s(\beta)$ is right continuous as a function of β.*

Before proving Lemma 4.3.3, consider the probability of survival until some fixed time t:

$$s(t, \beta) = P^O(\eta_t \neq \emptyset).$$

Most functions that depend only on the values of the process for a finite amount of time are continuous in β and $s(t, \beta)$ is no exception.

Proposition 4.3.4 *The function $s(t, \beta)$ is continuous in β for each fixed $t \ge 0$.*

Proof. By an argument similar to the one used to prove Proposition 2.4.1,

$$|\eta| \to |\eta| + 1 \qquad \text{at rate} \qquad \beta\left((d-1)|\eta| + 2\right)$$

provided $\eta \neq \emptyset$. Therefore, it is possible to couple the uniform model to a pure birth process Y_t that makes a transition from n to $n+1$ at rate $\beta\left((d-1)n + 2\right)$

such that

$$\left|\cup_{s\leq t}\eta_s\right| \leq Y_t.$$

In particular,

$$\mathbb{E}^O\left(\left|\cup_{s\leq t}\eta_s\right|\right) < \infty$$

for all $t \geq 0$.

Given $\beta < \beta^*$, the rates for the uniform model with these two parameters satisfy

$$c_\beta(x,\eta) \leq c_{\beta^*}(x,\zeta) \qquad \text{if } \zeta(x)=0 \qquad \text{and} \qquad c_\beta(x,\eta) \geq c_{\beta^*}(x,\zeta) \qquad \text{if } \eta(x)=1,$$

for $\eta \leq \zeta$. Therefore, it is possible to couple (η_t,ζ_t) two copies of the uniform model with parameters β and β^* respectively so that if $\eta_0 = \zeta_0$, then $\eta_t \subseteq \zeta_t$ for all $t \geq 0$. Let N_u be a Poisson process with a random parameter $(\beta^*-\beta)\left|\cup_{s\leq t}\eta_s\right|$. Thus,

$$\begin{aligned}P^{(O,O)}(\eta_t \neq \zeta_t \mid \eta_s, s\leq t) &\leq 1 - P(N_t = 0 \mid \eta_s, s \leq t)\\ &= 1-\exp\left(-(\beta^*-\beta)t\left|\cup_{s\leq t}\eta_s\right|\right).\end{aligned}$$

Taking expected value and using Jensen's inequality, one obtains

$$P^{(O,O)}(\eta_t \neq \zeta_t) \leq 1-\exp\left(-(\beta^*-\beta)t\mathbb{E}^O\left(\left|\cup_{s\leq t}\eta_s\right|\right)\right).$$

Letting β increase to β^* or β^* decrease to β has the consequence that $P^{(O,O)}(\eta_t \neq \zeta_t)$ tends to zero. In particular, $P^{(O,O)}(\eta_t = \emptyset, \zeta_t \neq \emptyset)$ tends to zero. ∎

The continuity of the probability of survival until time a fixed time t will be used in order to establish right continuity of the survival probability.

Proof of Lemma 4.3.3. Since $s(t, \beta)$ is continuous as a function of β, it follows that $s(t, \beta)$ is an upper semi-continuous function of β. Therefore,

$$P^O(\eta_t \neq \emptyset \ \forall \ t) = s(\beta) = \inf_t s(t, \beta)$$

is upper semi-continuous. By characterization i) of Proposition 4.3.2,

$$\limsup_{\beta \searrow \beta_0} s(\beta) \leq s(\beta_0). \tag{4.3.1}$$

Since $s(t, \beta)$ is decreasing as a function of t for each fixed β, it follows that $\lim_{t \to \infty} s(t, \beta)$ exists and is given by $s(\beta)$. Using the fact that $s(t, \beta)$ is increasing in β for each t and that the limit of increasing functions is increasing, it follows that $s(\beta)$ is an increasing as a function of β. Therefore, $\lim_{\beta \searrow \beta_0} s(\beta)$ exists and satisfies

$$\lim_{\beta \searrow \beta_0} s(\beta) \geq s(\beta_0).$$

Combining this with inequality (4.3.1), it follows that $\lim_{\beta \searrow \beta_0} s(\beta) = s(\beta_0)$. ∎

Lemma 4.3.5 *The function $s(\beta)$ is left continuous as a function of β on the interval in that the rooted chain is transient.*

Proof. Let

$$h(\beta, \hat{A}) = P^{\hat{A}}(\hat{A}_t \neq \hat{\emptyset} \ \forall \ t).$$

The function h is harmonic off $\hat{\emptyset}$ for the shape chain and it is immediate that $h(\beta,\hat{\emptyset}) = 0$. Furthermore, $h(\beta,\hat{O}) = s(\beta)$. Fix $A_n \subseteq \mathbb{T}^d$ finite and connected such that $O \in A_n$ and $|A_n| = n$. For $x_n \in A_n$, we can write A_n as the disjoint union $A_n \setminus x_n = \sqcup_{i=1}^{d+1} A_n^i$, where A_n^i are the $d+1$ connected components of $A_n \setminus x_n$. Choose $x_n \in A_n$ such that $|A_n^i|$ and $|A_n^j|$ tend to infinity for two distinct indices i and j. Without loss of generality, we may assume that $x_n = O$. By inequality (2.4.5),

$$h(\beta,\hat{A}_n) = P^{A_n}(\eta_t \neq \emptyset \ \forall\ t) \geq P(O \in \eta_t^{A_n} \text{ for all } t \geq 0)$$

$$\geq P(O \in \eta_t^{A_n} \text{ for all } t \leq u) P(S \leq u),$$

where S is defined as in that section. As was noted there,

$$P(\exists\ s \leq u \ni \eta_s^{A_n} \cap \mathbb{B}_i^d = \{O\}) \leq (1 - e^{-u})^{|A_n^i|}.$$

Therefore, the probability that this event occurs for at least one of the indices i or j tends to zero as n tends to infinity. In particular, $\liminf_{n\to\infty} h(\beta,\hat{A}_n) \geq P(S \leq u)$ for all $u > 0$. But the assumption that the rooted chain is transient implies that $P(S \leq u)$ tends to one as u tends to infinity. Therefore, $\lim_{n\to\infty} h(\beta,\hat{A}_n) = 1$, or equivalently $\lim_{|\hat{A}|\to\infty} h(\beta,\hat{A}) = 1$. Thus, $h(\beta,\cdot)$ is the unique harmonic function on $\hat{S}$ such that $h(\hat{\emptyset}) = 0$ and $\lim_{|\hat{A}|\to\infty} h(\beta,\hat{A}) = 1$.

Observe that $h(\beta,\hat{A})$ is increasing in β so that $\lim_{\beta\nearrow\beta^*} h(\beta,\hat{A})$ exists and is bounded by $h(\beta^*,\hat{A})$. Since $h(\beta,\cdot)$ is harmonic and since $\hat{q}(\hat{A},\hat{B}) > 0$ for finitely many $\hat{B}$ for any fixed $\hat{A}$, the limit $\lim_{\beta\nearrow\beta^*} h(\beta,\hat{A})$ is also harmonic. Furthermore,

$\lim_{\beta \nearrow \beta^*} h(\beta, \hat{\emptyset}) = 0$. Using the fact that $h(\beta, \hat{A})$ is increasing in β, it follows that

$$\lim_{|\hat{A}| \to \infty} \lim_{\beta \nearrow \beta^*} h(\beta, \hat{A}) = 1.$$

Therefore, $\lim_{\beta \nearrow \beta^*} h(\beta, \hat{A})$ is the unique harmonic function with these boundary conditions. Since $h(\beta^*, \hat{A})$ is also harmonic with the same boundary conditions, it follows that $\lim_{\beta \nearrow \beta^*} h(\beta, \hat{A}) = h(\beta^*, \hat{A})$. In particular, $\lim_{\beta \nearrow \beta^*} s(\beta) = s(\beta^*)$. ∎

Corollary 4.3.6 *On the binary tree, $s(\beta)$ is continuous.*

Proof. In the process of proving that $\beta_4(2) = 1/4$, it was shown that the rooted chain is transient for $\beta > 1/4$ (see Section 2.9). Therefore, Lemma 4.3.5 implies that $s(\beta)$ is left continuous for $\beta > 1/4$. By Theorem 1.4.1a), $s(\beta)$ is left continuous on $(0, 1/4]$ and therefore, it is everywhere left continuous. This together with Lemma 4.3.3 proves the assertion. ∎

4.4 Upper Bounds on the Survival Probability

This section is devoted to obtaining upper bounds on the survival probability via the Dirichlet principle.

Theorem 4.4.1 *For the shape chain with $\beta > \beta_1$,*

$$P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty) \leq C\beta \left(\frac{\beta - \beta_1}{\beta} \right)^{5/2} \tag{4.4.1}$$

for some constant $0 < C < \infty$.

Before proving Theorem 4.4.1, we pause to explain the origins of the functions that are used in the proof. We begin by fixing a sequence $\{\hat{\mathcal{S}}_N\}$ of subsets of $\hat{\mathcal{S}}$ that have the properties that $\hat{\mathcal{S}} \setminus \hat{\mathcal{S}}_N$ increases to $\hat{\mathcal{S}}$, $\hat{\emptyset} \notin \hat{\mathcal{S}}_N$, and $P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} < \infty) = 1$ for each $N \in \mathbb{N}$. Then, a function $h_N \in \mathcal{H}_{\hat{\emptyset}}^{\hat{\mathcal{S}}_N}$ is selected for each $N \in \mathbb{N}$. By the Dirichlet principle, $\beta P^{\hat{O}}(\tau_{\hat{\mathcal{S}}_N} < \tau_{\hat{\emptyset}}) \leq \Phi(h_N)$. Therefore,

$$\beta P^{\hat{\emptyset}}(\tau_{\hat{\emptyset}} = \infty) \leq \liminf_{N \to \infty} \Phi(h_N).$$

The idea is to choose h_N so that the liminf is as small as possible. Here h_N is chosen to be the minimizer of the Dirichlet form over all functions in $\mathcal{H}_{\hat{\emptyset}}^{\hat{\mathcal{S}}_N}$ that depend only on cardinality. In spite of the fact that the functions h_N take almost none of the structure of the sets into account, this choice of h_N provides a lower bound on β_2 that turns out to be equal to β_2 on the binary tree. It would be even more remarkable if such nondiscriminating functions provide the correct order of magnitude for the rate of decay of the survival probability. In fact, we expect that this is not the case and that the exponent given in bound (4.4.1) of Theorem 4.4.1 can be improved.

The next proposition is used in the proof of Theorem 4.4.1. It is an immediate consequence of Stirling's formula that says that $n! \sim n^n e^{-n}\sqrt{2\pi n}$, where $\sim$ means that the ratio tends to one.

Proposition 4.4.2 *For each $d \geq 2$, there exist constants $0 < K_1, K_2 < \infty$ such that*

$$\frac{K_2}{\sqrt{j}\beta_1^j} \leq \binom{dj}{j} \leq \frac{K_1}{\sqrt{j}\beta_1^j}$$

for each $j \geq 1$.

When it is necessary to emphasize which d is being considered, we write $K_1(d)$ (resp. $K_2(d)$) for K_1 (resp. K_2).

Proof of Theorem 4.4.1. Let $\hat{\mathcal{S}}_N = \{\hat{A} \in \hat{\mathcal{S}} : |\hat{A}| \geq N\}$. Also, let $g_N : \{1, \ldots, N\} \to \mathbb{R}$ be given by

$$g_N(k+1) = \begin{cases} 1 & \text{if } k = 0, \\ g_N(k) + \frac{k}{(k+2)t(k)\beta^k} & \text{otherwise,} \end{cases}$$

where $t(k)$ is the number of connected subsets of $\mathbb{T}^d$ of size k containing O. Also, define

$$h_N(\hat{A}) = \begin{cases} 0 & \text{if } \hat{A} = \hat{\emptyset}, \\ \frac{g_N(|\hat{A}|)}{g_N(N)} & \text{if } 1 \leq |\hat{A}| \leq N, \\ 1 & \text{otherwise.} \end{cases}$$

Note that $h_N \in \mathcal{H}_{\hat{\emptyset}}^{\hat{\mathcal{S}}_N}$. Let $\mathcal{N}(\hat{A}) = \{\hat{B} : \hat{q}(\hat{A}, \hat{B}) > 0, |\hat{A}| < |\hat{B}|\}$. Using the fact that $\sum_{\hat{B} \in \mathcal{N}(\hat{A})} |\{B \in \hat{B} : A \subset B\}| = |A| + 2$ and $\sum_{|\hat{A}|=n} M(\hat{A}) = t(n)$,

$$\begin{aligned}
\Phi(h_N) &= \frac{\beta}{g_N^2(N)} + \sum_{n=1}^{N-1} \sum_{|\hat{A}|=n} \sum_{\hat{B} \in \mathcal{N}(\hat{A})} \frac{M(\hat{A})\beta^{n+1}}{n} |\{B \in \hat{B} : A \subset B\}| \left(\frac{\frac{n}{(n+2)t(n)\beta^n}}{g_N(N)} \right)^2 \\
&= \frac{\beta}{g_N^2(N)} \left(1 + \sum_{n=1}^{N-1} \frac{t(n)(n+2)}{n} \beta^n \left(\frac{n}{(n+2)t(n)\beta^n} \right)^2 \right) \\
&= \frac{\beta}{g_N(N)}.
\end{aligned}$$

By the Dirichlet principle,

$$P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} \le \tau_{\hat{\emptyset}}^+) \le \frac{1}{g_N(N)}.$$

Since $P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} \le \tau_{\hat{\emptyset}}^+) = P^{\hat{O}}(\tau_{\hat{\mathcal{S}}_N} \le \tau_{\hat{\emptyset}})$, it follows that

$$P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty) \le \liminf_{N \to \infty} \frac{1}{g_N(N)} = \frac{1}{1 + \sum_{n=1}^{\infty} \frac{n}{(n+2)t(n)\beta^n}} \le \frac{1}{\sum_{n=1}^{\infty} \frac{n}{(n+2)t(n)\beta^n}}$$

Recall that $c(n)$, the number of connected subsets of $\mathbb{B}^d$ containing the root, is given by $\binom{dn}{n}/((d-1)n+1)$ (see equation (2.5.6)). The quantities $c(n)$ and $t(n)$ are related via the recursion

$$t(n) = \sum_{k_1 + \cdots + k_{d+1} = n-1} c(k_1) \cdots c(k_{d+1}) \qquad \text{for } n \ge 1. \tag{4.4.2}$$

This follows from an argument similar to the one use to prove recursion (2.5.2). By requiring $k_1 + \cdots + k_d = k$ and applying recursion (2.5.2),

$$t(n) = \sum_{k=0}^{n-1} c(k+1)c(n-1-k). \tag{4.4.3}$$

Therefore, $t(n) \leq c(n+1)$. Combining this with the definition of $c(j)$ and Proposition 4.4.2 gives

$$(j+2)t(j) \leq (j+2)c(j+1) \leq \binom{d(j+1)}{j+1} \leq \frac{K_1(d)}{\sqrt{j+1}\beta_1^{j+1}}.$$

Therefore,

$$P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty) \leq \frac{K_1(d)}{\beta_1 \sum_{j=1}^{\infty} j\sqrt{j+1}\left(\frac{\beta_1}{\beta}\right)^j}.$$

Making the transformation $s = \beta_1/\beta$, it suffices to obtain an appropriate lower bound on the series

$$\sum_{j=0}^{\infty}(j+1)\sqrt{j+2}\ s^j$$

for $0 \leq s < 1$. By expanding in a power series about zero,

$$\frac{1}{(1-s)^{\frac{5}{2}}} = \frac{2}{3}\sum_{k=0}^{\infty}\frac{(2k+3)(k+1)}{4^{k+1}}\binom{2(k+1)}{(k+1)}s^k.$$

Using Proposition 4.4.2,

$$\frac{(2n+3)(n+1)}{4^{n+1}}\binom{2(n+1)}{n+1} \leq K_1(2)(2n+3)\sqrt{n+1} \leq K_1(2)3(n+1)\sqrt{n+2}$$

for $n \geq 0$. Therefore, it follows that

$$P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty) \leq \frac{2K_1(2)K_1(d)\beta}{\beta_1^2}\left(\frac{\beta-\beta_1}{\beta}\right)^{5/2}. \qquad \blacksquare$$

Corollary 4.4.3 *The survival probability satisfies*

$$\limsup_{\beta \searrow \beta_1} \frac{P(\eta_t^O \neq \emptyset\ \forall\ t)}{(\beta-\beta_1)^{\frac{5}{2}}} \leq C_1$$

for some $0 < C_1 < \infty$. *In particular, the critical exponent (if it exists) is greater than one.*

Proof. The uniform model and the shape chain $\hat{A}_t^{\hat{O}}$ maybe coupled such that $\eta_t^O \in \hat{A}_t^{\hat{O}}$ for all $0 \le t \le \tau_{\hat{\emptyset}}$. Therefore, $P(\eta_t^O \neq \emptyset \ \forall \ t) = P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty)$ so that both statements follow directly from Theorem 4.4.1. ∎

4.5 Lower Bounds on the Survival Probability in $d = 2$

This section is devoted to obtaining lower bounds on the survival probability via Thompson's principle. In order to do this, observe that a function $\hat{w} : \hat{\mathcal{S}} \times \hat{\mathcal{S}} \to \mathbb{R}$ is an *anti-symmetric, incompressible flow* on $\hat{\mathcal{S}}$ according to Definition 2.7.1 if $\hat{w} \in \mathcal{W}_{\hat{\emptyset}}^{\hat{\mathcal{S}}_N}$ for every $N \in \mathbb{N}$. By Thompson's Principle,

$$\frac{1}{\mathcal{K}(\hat{w})} \le \sup_{\hat{w} \in \mathcal{W}_{\hat{\emptyset}}^{\hat{\mathcal{S}}_N}} \frac{1}{\mathcal{K}(w)} = \beta P^{\hat{\emptyset}}(\tau_{\hat{\mathcal{S}}_N} < \tau_{\hat{\emptyset}}^+) = \beta P^{\hat{O}}(\tau_{\hat{\mathcal{S}}_N} < \tau_{\hat{\emptyset}}).$$

By letting N tend to infinity,

$$\frac{1}{\mathcal{K}(\hat{w})} \le \beta P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty).$$

Therefore, in order to obtain lower bounds on the survival probability, it suffices to construct a flow on $\hat{\mathcal{S}}$ and estimate the energy.

The flow analyzed here was construct in Section 2.9 for the rooted chain. That flow is lifted to the shape chain providing lower bounds on the probability of survival and upper bounds on the critical exponent. Unfortunately, that flow was only completely constructed for $d = 2$ which explains the specialization to the binary tree in this section. The contribution here is the estimate on the energy.

Presumably, the techniques used to estimate the energy can be executed more generally provided that the flow can be constructed. This is discussed more fully at the end of this section.

Before proceeding to define the lift, we review the definition of the rooted chain and the flow that was constructed in Section 2.9. Let $\{x_1, x_2, x_3\}$ denote the 3 vertices adjacent to the root O. Let $\mathbb{B}^2 = \{x \in \mathbb{T}^2 : \|x - x_1\| \leq \|x - O\|\} \cup O$. Consider the initial configuration $\eta_0 = \mathbb{T}^2 \setminus \mathbb{B}^2 \cup O$. By connectedness, $\eta_0 \subseteq \eta_t$ for all $t \geq 0$. Let $A_t = \mathbb{B}^2 \cap \eta_t$ so that A_t is a Markov chain with state space $\mathcal{C}_2 = \{\text{finite, connected } A \subset \mathbb{B}^2 \text{ containing } O\}$. Recall that A_t is reversible with stationary measure $\pi(A) = \beta^{|A|}$, where $|A|$ is number of vertices in $A \setminus O$. Since $O \in A_t$ for all t, we refer to A_t as the rooted chain on $\mathbb{B}^2$.

The advantage of constructing flows for the rooted chain is that its state space allows flows to be constructed recursively by taking advantage of self similarity properties of $\mathbb{B}^2$: see Section 2.7. The *uniformly distributed flow* is defined in the following manner. For each $n \geq 1$, let

$$\alpha(n, k) = \frac{(k+1)(2k+1)(3n-2k)}{n(n+1)(2n+1)} \tag{4.5.1}$$

for $0 \leq k \leq n-1$. Note that $\alpha(n, k) \geq 0$ and that $\alpha(n, k) + \alpha(n, n-1-k) = 1$. Given a set $A \in \mathcal{C}_2$, let $A_i = A \cap \mathbb{B}^2_{1i}$, $i = 1, 2$. Here $\{y_1, y_2\}$ are the neighbors of x_1 in $\mathbb{B}^d$, neither of which is O, and $\mathbb{B}^2_{1i} = \{y : \|y_i - y\| \leq \|x_1 - y\|\} \cup x_1$. For each

$A \in \mathcal{C}_2$, define the map $r(A, \cdot)$ with domain $\mathcal{N}_2(A) = \{B : q(A,B) = \beta\}$ by

$$r(A,B) = \begin{cases} 1 & \text{if } (A,B) = (\emptyset, O), \\ \alpha(|A|, |A_i|) r(A_i, B_i) & \text{if } A \neq \emptyset \text{ and } B_i \in \mathcal{N}_2(A_i), \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $r(A, \cdot) \geq 0$ and $\sum_{B \in \mathcal{N}_2(A)} r(A,B) = 1$ for each A. Finally,

$$w(A,B) = \begin{cases} r(A,B)/c(n) & \text{if } |A| = n \text{ and } B \in \mathcal{N}_2(A), \\ -w(B,A) & \text{if } A \in \mathcal{N}_2(B), \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 2.9.2 and equation (2.9.5),

$$\sum_{\{B : A \in \mathcal{N}_2(B)\}} w(B,A) = \frac{1}{c(n)}$$

for each $A \in \mathcal{C}_2(n)$ and for each $n \in \mathbb{N}$. This implies that w satisfies the incompressibility property and justifies calling w the uniformly distributed flow on $\mathcal{C}_2$.

The next objective is to lift the uniformly distributed flow on $\mathcal{C}_2$ to the state space of the shape chain $\hat{\mathcal{S}}$. For this purpose define an anti-symmetric function on $\hat{\mathcal{S}} \times \hat{\mathcal{S}}$ by

$$\hat{w}(\hat{A}, \hat{B}) = \sum_{\{A \in \hat{A} : x_1 \in A, O \notin A\}} \sum_{\{B \in \hat{B} : x_1 \in B, O \notin B\}} w(O \cup A, O \cup B) \tag{4.5.2}$$

for $\hat{A}, \hat{B} \neq \hat{\emptyset}$. Also, set $\hat{w}(\hat{\emptyset}, \hat{O}) = -\hat{w}(\hat{O}, \hat{\emptyset}) = 1$. It is immediate that $\hat{w}$ is an anti-symmetric, incompressible flow. The energy of this flow is given by

$$\begin{aligned}\mathcal{K}(\hat{w}) &= \sum_{n=0}^{\infty} \sum_{|\hat{A}|=n} \sum_{\hat{B}\in\mathcal{N}_2(\hat{A})} \frac{\hat{w}^2(\hat{A},\hat{B})}{\hat{\pi}(\hat{A})\hat{q}(\hat{A},\hat{B})} \\ &= \frac{1}{\beta} + \sum_{n=1}^{\infty} \frac{n}{\beta^{n+1}} \sum_{|\hat{A}|=n} \sum_{\hat{B}\in\mathcal{N}_2(\hat{A})} \frac{\left(\sum_{\{A\in\hat{A}:x_1\in A, O\notin A\}} \sum_{\{B\in\hat{B}:x_1\in B, O\notin B\}} w(O\cup A, O\cup B)\right)^2}{M(\hat{A})|\{B\in\hat{B}: A\subset B\}|}.\end{aligned}$$

By the Cauchy-Schwarz inequality and the fact that for each pair $\hat{A}$ and $\hat{B}$, the number of terms that appears in the numerator is at most $M(\hat{A})|\{B\in\hat{B}: A\subset B\}|$,

$$\begin{aligned}\mathcal{K}(\hat{w}) &\le \frac{1}{\beta} + \sum_{n=1}^{\infty} \frac{n}{\beta^{n+1}} \sum_{|\hat{A}|=n} \sum_{\hat{B}\in\mathcal{N}_2(\hat{A})} \sum_{\{A\in\hat{A}:x_1\in A, O\notin A\}} \sum_{\{B\in\hat{B}:x_1\in B, O\notin B\}} w^2(O\cup A, O\cup B) \\ &= \frac{1}{\beta} + \sum_{n=1}^{\infty} \frac{n}{\beta^{n+1}} \sum_{A\in\mathcal{C}_2(n)} \sum_{B\in\mathcal{N}_2(A)} w^2(A,B). \qquad (4.5.3)\end{aligned}$$

Recall that $\mathcal{C}_2(n) = \{A\in\mathcal{C}_2 : |A| = n\}$. Using the definition of w and inequality (4.5.3),

$$\mathcal{K}(\hat{w}) \le \frac{1}{\beta} + \sum_{n=1}^{\infty} \frac{n}{c(n)^2\beta^{n+1}} \sum_{A\in\mathcal{C}_2(n)} \sum_{B\in\mathcal{N}_2(A)} r^2(A,B) \qquad (4.5.4)$$

Since the asymptotic behavior of $c(n)$ is known, it would suffice to determine the asymptotic behavior of

$$g(n) = \sum_{A\in\mathcal{C}_2(n)} \sum_{B\in\mathcal{N}_2(A)} r^2(A,B).$$

However, it turns out to be more manageable to determine the asymptotic behavior of the series itself. By Proposition 4.4.2,

$$\sum_{n=1}^{\infty} \frac{ng(n)}{c(n)^2\beta^{n+1}} \le \frac{1}{K_2^2(2)\beta} \sum_{n=1}^{\infty} \frac{g(n)n^2(n+1)^2}{(16\beta)^n}. \qquad (4.5.5)$$

Making the substitution $s = 1/16\beta$, the series of interest becomes

$$\sum_{n=1}^{\infty} g(n)n^2(n+1)^2 s^n \qquad (4.5.6)$$

as s increases to $1/4$.

We begin by showing that $g(n)$ satisfies a certain recursion. This recursion implies that a series similar to series (4.5.6) is a solution of an ordinary differential equation. The ordinary differential equation takes a particularly nice form. In fact, fairly elementary techniques allow one to exhibit the general solution of the ordinary differential equation. This provides an alternative representation of the series solution. This alternative representation readily reveals the asymptotic behavior of the series solution as s increases to $1/4$. Relating the series solution of the ordinary differential equation to series (4.5.6) gives a lower bound on the survival probability that implies inequality (1.4.2) in Theorem 1.4.4.

Proposition 4.5.1 *For* $n \geq 1$,

$$g(n) = 2 \sum_{k=0}^{n-1} c(n-1-k) g(k) \alpha^2(n,k).$$

Proof. By definition of $g(n)$ and $r(A, B)$,

$$\begin{aligned} g(n) &= \sum_{|A|=n} \sum_{B \in \mathcal{N}_2(A)} r^2(A,B) \\ &= \sum_{|A|=n} \left(\alpha^2(n, |A_1|) \sum_{B_1 \in \mathcal{N}_2(A_1)} r^2(A_1, B_1) + \alpha^2(n, |A_2|) \sum_{B_2 \in \mathcal{N}_2(A_2)} r^2(A_2, B_2) \right) \end{aligned}$$

for $n \geq 1$. By conditioning on the size of A_1,

$$\begin{aligned} g(n) =& \sum_{k=0}^{n-1} \sum_{\{|A|=n, |A_1|=k\}} \alpha^2(n,k) \sum_{B_1 \in \mathcal{N}_2(A_1)} r^2(A_1, B_1) \\ &+ \sum_{k=0}^{n-1} \sum_{\{|A|=n, |A_1|=k\}} \alpha^2(n, n-1-k) \sum_{B_2 \in \mathcal{N}_2(A_2)} r^2(A_2, B_2) \\ =& 2\sum_{k=0}^{n-1} \sum_{\{|A|=n, |A_1|=k\}} \alpha^2(n,k) \sum_{B_1 \in \mathcal{N}_2(A_1)} r^2(A_1, B_1). \end{aligned}$$

Using the fact that $\alpha^2(n,k) \sum_{B_1 \in \mathcal{N}_2(A_1)} r^2(A_1, B_1)$ is independent of A_2,

$$\begin{aligned} g(n) =& 2\sum_{k=0}^{n-1} \sum_{\{|A_1|=k\}} c(n-1-k)\alpha^2(n,k) \sum_{B_1 \in \mathcal{N}_2(A_1)} r^2(A_1, B_1) \\ =& 2\sum_{k=0}^{n-1} c(n-1-k)\alpha^2(n,k) g(k). \qquad \blacksquare \end{aligned}$$

As a consequence of Proposition 4.5.1, we obtain the next lemma.

Lemma 4.5.2 *Let* $G(s) = \sum_{n=0}^{\infty} g(n)(n+1)^2(2n+1)^2 s^n$. *Then* $G(0) = 1$, $G(s)$ *converges for* $|s| < 1/4$, *and* $G(s)$ *is a solution of*

$$s(1-4s)^2 G''(s) + (1-16s)(1-4s)G'(s) - 18(1-2s)G(s) = 0 \qquad (4.5.7)$$

for $|s| < 1/4$.

Proof. By Cauchy-Schwarz,

$$\frac{1}{n+1} \leq \sum_{B \in \mathcal{N}_2(A)} r^2(A, B) \leq 1 \qquad \text{for } |A| = n.$$

Thus,

$$\frac{c(n)}{n+1} \leq g(n) \leq c(n)$$

This together with Proposition 4.4.2 implies that the series defining $G(s)$ converges for $|s| < 1/4$.

By Proposition 4.5.1 and equation (4.5.1),

$$g(n)(n+1)^2(2n+1)^2n^2 = 2\sum_{k=0}^{n-1} c(n-1-k)g(k)(k+1)^2(2k+1)^2(3(n-k)+k)^2$$

for $n \geq 1$. Expanding the factor $(3(n-k)+k)^2$ as $9(n-k)^2 + 6(n-k)k + k^2$, multiplying by s^{n-1}, and taking the sum from $n = 1$ to infinity, implies that

$$\begin{aligned} G'(s) + sG''(s) =& 18\left(C(s) + 3sC'(s) + s^2C''(s)\right)G(s) \\ &+ 12\left(C(s) + sC'(s)\right)sG'(s) + 2C(s)\left(sG'(s) + s^2G''(s)\right) \end{aligned}$$

where $C(s) = \sum_{n=0}^{\infty} c(n)s^n$. Multiplying recursion (2.5.2) by s^{n-1} and taking the sum from $n = 1$ to infinity gives

$$C(s) = \frac{1-\sqrt{1-4s}}{2s} \qquad \text{for } 0 \leq s \leq 1/4.$$

Using the explicit expression for $C(s)$,

$$s(1-4s)^2G''(s) + (1-16s)(1-4s)G'(s) - 18(1-2s)G(s) = 0 \qquad G(0) = 1. \qquad \blacksquare$$

We will show that the ordinary differential equation determines the rate at which $G(s)$ tends to infinity as s increases to $1/4$. Since the coefficient of $G''(s)$ in the ordinary differential equation has a factor of s, it follows that there are solutions to the ordinary differential equation that also blow up as s tends to zero. On $(0, 1/4)$, the general solution to the ordinary differential equation is of the form

$$c_1(1-4s)^{r_1}H_1(s) + c_2(1-4s)^{r_2}H_2(s) \tag{4.5.8}$$

where c_i, $i = 1, 2$, are arbitrary constants, $r_1 = -1 - \sqrt{13}/2$, $r_2 = -1 + \sqrt{13}/2$, and $H_i(s) = \sum_{n=0}^{\infty} h_i(n)(1-4s)^n$ with $h_i(0) = 1$, $i = 1, 2$, and

$$h_i(n) = \frac{2n^2 + (2 + 4r_i)n + 5 - 2r_i}{2n^2 + (4 + 4r_i)n} h_i(n-1), \qquad (4.5.9)$$

for $n \geq 1$.

In order to see that expression (4.5.8) is the general solution, let

$$H(s) = \sum_{n=0}^{\infty} h(n)(1-4s)^{n+r},$$

where $h(n)$ is defined as in equation (4.5.9) except with r_i replaced by r. We have

$$\begin{aligned} H(s) &= \sum_{n=0}^{\infty} h(n)(1-4s)^{n+r} \\ (1-4s)H'(s) &= -4\sum_{n=0}^{\infty} h(n)(n+r)(1-4s)^{n+r} \\ (1-4s)^2 H''(s) &= 16\sum_{n=0}^{\infty} h(n)(n+r)(n+r-1)(1-4s)^{n+r}. \end{aligned}$$

Also, expressing the coefficients of $(1-4s)^n G^{(n)}(s)$ in equation (4.5.7) as linear combinations of $\{(1-4s)^m\}_{m \in \mathbb{N}}$ gives

$$-18(1-2s) = -9 - 9(1-4s), \qquad (1-16s) = -3 + 4(1-4s),$$

$$\text{and} \qquad s = 1/4 - (1-4s)/4.$$

Therefore,

$$
\begin{aligned}
-18(1-2s)H(s) =& -9h(0) - 9\sum_{n=1}^{\infty}(h(n)+h(n-1))(1-4s)^{n+r}, \\
(1-16s)(1-4s)H'(s) =& 12h(0)r + \sum_{n=1}^{\infty}(12h(n)(n+r) \\
& - 16h(n-1)(n+r-1))(1-4s)^{n+r}, \quad \text{and} \\
s(1-4s)^2H''(s) =& 4h(0)r(r-1) + 4\sum_{n=1}^{\infty}(h(n)(n+r)(n+r-1) \\
& - h(n-1)(n+r-1)(n+r-2))(1-4s)^{n+r}.
\end{aligned}
$$

Adding these three expressions and combining like terms shows that $H(s)$ is a solution if and only if

$$
\begin{aligned}
h(0)(-9+8r+4r^2) &= 0 \\
h(n)(4n^2+(8+8r)n) &= h(n-1)(4n^2+(4+8r)n+10-4r).
\end{aligned}
$$

In particular, $(1-4s)^{r_i}H_i(s)$, $i=1,2$, are two linearly independent solutions to the ordinary differential equation.

Since all solutions on $(0,1/4)$ are given by expression (4.5.8), there exists a choice of c_i, $i=1,2$, such that

$$G(s) = c_1(1-4s)^{r_1}H_1(s) + c_2(1-4s)^{r_2}H_2(s)$$

on $(0,1/4)$. The fact that $r_2>0$ implies that $c_2(1-4s)^{r_2}H_2(s)$ tends to zero as s increases to $1/4$. Since $G(s)$ tends to infinity as s increases to $1/4$, it follows that $c_1>0$ and

$$G(s) \sim c_1(1-4s)^{r_1} \qquad \text{as } s \nearrow 1/4.$$

Equivalently,

$$G(1/16\beta) \sim c_1 \left(\frac{\beta - 1/4}{\beta}\right)^{r_1} \qquad \text{as } \beta \searrow 1/4. \tag{4.5.10}$$

Theorem 4.5.3 *For the shape chain on the binary tree,*

$$C_2 \le \liminf_{\beta \searrow 1/4} \frac{P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty)}{(\beta - 1/4)^{1+\sqrt{13}/2}} \tag{4.5.11}$$

for some constant $0 < C_2 < \infty$.

Proof. By Thompson's principle, inequalities (4.5.4) and (4.5.5), and the fact that $n \le (2n+1)/2$,

$$\frac{1}{1 + \frac{1}{4K_2^2(2)\beta} G(1/16\beta)} \le \frac{1}{1 + \sum_{n=1}^{\infty} \frac{ng(n)}{c^2(n)\beta^n}} \le P^{\hat{O}}(\tau_{\emptyset} = \infty).$$

By asymptotic relation (4.5.10),

$$\frac{K_2^2(2)}{c_1 4^{1+\sqrt{13}/2}} \le \liminf_{\beta \searrow 1/4} \frac{P^{\hat{O}}(\tau_{\hat{\emptyset}} = \infty)}{(\beta - 1/4)^{1+\sqrt{13}/2}}. \qquad \blacksquare$$

Corollary 4.4.3 and Theorem 4.5.3 imply that if the critical exponent exists, then on the binary tree it is lies in the interval $[5/2, 1 + \sqrt{13}/2]$. If one could show that the hypothesis of Lemma 2.9.1 holds for $d \ge 3$, then the analog of $g(n)$ satisfies $g(n) \le c(n)$ which gives an upper bound of $7/2$ on the critical exponent on the d-ary tree. Furthermore, the analog of the recursion in Proposition 4.5.1 will hold:

$$g(n) = d \sum_{k=0}^{n-1} \sum_{j_1 + \cdot + j_{d-1} = n-1-k} c(j_1) \cdots c(j_{d-1}) \alpha^2(n; (k, j_1, \ldots, j_{d-1})) g(k).$$

If one provides more information about $\alpha(n;\cdot)$ in proving that the hypothesis of Lemma 2.9.1 holds, then it may be possible to improve the exponent to $1+\sqrt{13}/2$. It is unclear whether or not either of these bounds provides a sharp estimate on the rate of decay.

4.6 The Expected Extinction Time and the Susceptibility

Sections 4.4 and 4.5 bounded the rate at which the survival probability approaches zero as β decreases to β_1. In the subcritical regime, there are other quantities that typically diverge as β increases to β_1. For example, the expected extinction time is often infinite at the critical point. In the case of the uniform model, such quantities do not diverge because the shape chain exhibits positive recurrent behavior at β_1. Instead they approach some constant. Due to reversibility, more information than simply the rate at which these quantities approach some constant can be provided. In fact, explicit formulas for the expected extinction time and the susceptibility are obtainable. The derivations of these formulas are given here.

Theorem 4.6.1 *For $\beta \le \beta_1$,*

$$a)\ \mathbb{E}^O(\tau) = \frac{1}{\beta}\int_0^\beta \frac{C^2(x)-C(x)}{x}dx \quad and \quad b)\ \mathbb{E}\int_0^\infty |\eta_t^O|\,dt = \frac{C^2(\beta)-C(\beta)}{\beta}.$$

Proof. Recall that $t(n)$ is the number of connected subsets of $\mathbb{T}^d$ that contain O and $c(n)$ is the number of connected subsets of $\mathbb{B}^d$ that contain O. The sequences $t(n)$ and $c(n)$ are related to each other via recursion (4.4.3). Therefore,

$$t(n) = \sum_{k=0}^{n} c(k)c(n-k) - c(n), \tag{4.6.1}$$

for $n \geq 1$. The normalizing constant for the stationary measure of the shape chain is given by

$$\hat{C}(\beta) = 1 + \sum_{n=1}^{\infty} \frac{t(n)}{n} \beta^n.$$

By recursion (4.6.1),

$$\hat{C}'(\beta) = \frac{C^2(\beta) - C(\beta)}{\beta}. \tag{4.6.2}$$

Since the shape chain is positive recurrent,

$$\mathbb{E}^{\hat{\emptyset}} \int_0^{\tau_{\hat{\emptyset}}^+} h(\hat{A}_t)\mathrm{d}t = \frac{\sum_{\hat{A}} h(\hat{A})\hat{\pi}(\hat{A})}{\beta}$$

for any nonnegative function h on $\hat{S}$. Taking $h \equiv 1$ gives

$$\mathbb{E}^{\hat{\emptyset}}(\tau_{\hat{\emptyset}}^+) = \frac{\hat{C}(\beta)}{\beta}.$$

By expressing $\tau_{\hat{\emptyset}}^+$ as the sum of the time to hit $\hat{O}$ starting from $\hat{\emptyset}$ and the time to hit the $\hat{\emptyset}$ starting from $\hat{O}$,

$$\mathbb{E}^{\hat{\emptyset}}(\tau_{\hat{\emptyset}}^+) = \frac{1}{\beta} + \mathbb{E}^{\hat{O}}(\tau_{\hat{\emptyset}}).$$

Combining these two expressions and using equation (4.6.2),

$$\mathbb{E}^{\hat{O}}(\tau_{\hat{\emptyset}}) = \frac{\hat{C}(\beta) - 1}{\beta} = \frac{1}{\beta} \int_0^{\beta} \frac{C^2(x) - C(x)}{x} \mathrm{d}x$$

which proves a). Taking $h(\hat{A}) = |\hat{A}|$ and using equation (4.6.2) gives

$$\mathbb{E}^{\hat{\emptyset}} \int_0^{\tau_{\hat{\emptyset}}^+} |\hat{A}_t| \mathrm{d}t = \sum_{n=1}^{\infty} t(n) \beta^{n-1} = \hat{C}'(\beta) = \frac{C^2(\beta) - C(\beta)}{\beta}$$

which proves b). ■

On the binary tree,

$$C(\beta) = \frac{1 - \sqrt{1 - 4\beta}}{2\beta}$$

so that

$$\hat{C}'(\beta) = \frac{1 - 3\beta - (1 - \beta)\sqrt{1 - 4\beta}}{2\beta^3}.$$

Integrating this expression and using the fact that $C(0) = 1$,

$$\hat{C}(\beta) = \frac{-1 + 6\beta + (1 - 4\beta)^{3/2}}{4\beta^2} - \frac{1}{2}.$$

This implies that

$$\mathbb{E}^O(\tau) = \frac{-1 + 6\beta - 6\beta^2 + (1 - 4\beta)^{3/2}}{4\beta^3}$$

$$\mathbb{E} \int_0^{\tau} |\eta_t^O| \mathrm{d}t = \frac{1 - 3\beta - (1 - \beta)\sqrt{1 - 4\beta}}{2\beta^3},$$

when $d = 2$. In particular,

$$\lim_{\beta \nearrow 1/4} \frac{2 - \mathbb{E}^O(\tau)}{1/4 - \beta} = 6 \quad \text{and} \quad \lim_{\beta \nearrow 1/4} \frac{8 - \mathbb{E} \int_0^{\tau} |\eta_t^O| \mathrm{d}t}{\sqrt{1/4 - \beta}} = 24.$$

CHAPTER 5

Reversible Invariant Measures

We now turn our attention to the infinite system and the study of invariant measures. By now, it is not hard to see that there are exactly two invariant measures in the supercritical regime, the pointmass on the empty configuration and the pointmass on the completely occupied configuration. This follows from the fact that the finite system obeys complete convergence. A more interesting question arises when one asks about the subcritical regime. Unlike the contact process and one dimensional reversible nearest particle systems, the uniform model has a rather large collection of subcritical invariant measures. Here, that collection of reversible invariant measures is described.

5.1 Some Background on Reversible Measures

A probability measure $\mu \in \mathcal{P}$ is said to be invariant for the process if $\mu = \mu S(t)$ for all $t \geq 0$. The set of all invariant measures is denoted by $\mathcal{J}$. Some basic properties of invariant measures are collected in the next proposition. The proofs of these statements can be found in Liggett [17].

Proposition 5.1.1

a) $\mu \in \mathcal{J}$ *if and only if*

$$\int S(t)f\,d\mu = \int f\,d\mu$$

for all $f \in C(X)$ *and for all* $t \geq 0$.

b) $\mathcal{J}$ *is a nonempty, compact convex subset of* $\mathcal{P}$.

c) Let $\mathcal{J}_e$ *be the extreme points of* $\mathcal{J}$. *Then* $\mathcal{J}$ *is the closed convex hull of* $\mathcal{J}_e$.

d) If $\nu = \lim_{t\to\infty} S(t)d\mu$, *then* $\nu \in \mathcal{J}$.

Definition 5.1.2 *A probability measure* $\mu \in \mathcal{P}$ *is said to be reversible if*

$$\int fS(t)g\,d\mu = \int gS(t)f\,d\mu, \tag{5.1.1}$$

for all $f,\ g \in C(X)$. *In other words,* $S(t)$ *is self-adjoint with respect to* μ.

The next theorem provides a useful characterization of reversibility in the context of spin systems.

Theorem 5.1.3 *Let* $c(x,\eta)$ *be the rates for a spin system and let* μ *be probability measure on* X. *The measure* μ *is reversible if and only if*

$$\int_X c(x,\eta)f(\eta)\,d\mu(\eta) = \int_X c(x,\eta)f(\eta_x)\,d\mu(\eta)$$

for every cylinder function f *and for every* $x \in \mathbb{T}^d$.

5.2 Subcritical Reversible Measures

Say that $b \subseteq \mathbb{T}^d$ is a *backbone* if b is nonempty and has no leaves. Let $\partial b = \{x \in \mathbb{T}^d : \|x - b\| = 1\}$ be the *exterior boundary* of b. Also, for $x \in \partial b$, let x^* denote the unique element of b at distance one from x. For $x \in \partial b$, define $B_b(x) = \{y : \|x - y\| \le \|b - y\|\} \cup x^*$ to be the branch of $\mathbb{T}^d$ extending from b into b^c through x. Also, let $\mathcal{C}_b(x)$ be the set of all finite, connected subsets of $B_b(x)$ containing x together with the empty set. Finally, let $\pi_{b,x}$ be the probability measure on $\mathcal{C}_b(x)$ given by

$$\pi_{b,x}(A) = \frac{\beta^{|A|}}{C(\beta)}.$$

Here $C(\beta)$ is a normalizing constant that depends on d as well as β. The assumption that $\beta \le \beta_1$ implies that such a normalizing constant exists.

The measures $\{\pi_{b,x}\}_{x \in \partial b}$ induce a measure μ_b on X that is given by

$$\mu_b = \mathbf{1}_{\{b \subseteq \eta\}} \prod_{x \in \partial b} \pi_{b,x}.$$

In particular, the support of μ_b is the set $X_b = \{\eta : b \subseteq \eta \text{ and } A_x^b(\eta) \in \mathcal{C}_b(x)\ \forall\ x \in \partial b\}$, where $A_x^b(\eta) = \eta \cap B_b(x)$. Since $\mu_b \perp \mu_{b'}$ for all distinct backbones b and b', it follows that the mapping from backbones to probability measures is one-to-one.

Proposition 5.2.1 *$\{\mu_b : b$ is a backbone$\} \subseteq \mathcal{R}_e$.*

Proof. Fix a backbone b and let $(A_t^{b,x})_{x \in \partial b}$ be a collection of independent rooted chains such that for each $x \in \partial b$ the state space of $A_t^{b,x}$ is $C_b(x)$ and the initial state $A_0^{b,x}$ is $\emptyset$. Set

$$\eta_t^b = b \cup \{ \bigcup_{x \in \partial b} A_t^{b,x} \}.$$

Thus, η_t^b is distributed as a uniform model with initial state b such that $\eta_t^b \cap B_b(x) = A_t^{b,x}$ for all $t \geq 0$. Since $A_t^{b,x}$ is converging in distribution to $\pi_{b,x}$, and since any cylinder function depends only on the state of finitely many rooted chains, it follows that η_t^b converges in distribution to μ_b. Therefore, μ_b is invariant.

Suppose that

$$\mu_b = \lambda \mu_1 + (1 - \lambda) \mu_2 \tag{5.2.1}$$

for some measures $\mu_i \in \mathcal{J}$. In this case, $\mu_i(\eta : b \subseteq \eta) = 1$ for $i = 1, 2$ so that $\delta_b \leq \mu_i$. Since the uniform model is an attractive spin system, $\delta_b S(t) \leq \mu_i S(t) = \mu_i$ and consequently $\mu_b \leq \mu_i$. By equation (5.2.1), $\mu_b = \mu_i$, and consequently μ_b is extremal.

Finally, for any pair of cylinder functions f and g, there exists a finite collection $\{x_i\}_{i=0}^n$ of vertices in ∂b such that f and g are determined by the states of the associated rooted chains $\{A_t^{b,x_i}\}_{i=0}^n$. Without loss of generality, write

$$f(\eta) = f(A_{x_0}^b(\eta), \ldots, A_{x_n}^b(\eta)),$$

where $A_x^b(\eta) = \eta \cap B_b(x)$ as before. By definition,

$$\begin{aligned} &\int_X g(\eta)\ S(t) f(\eta)\ \mathrm{d}\mu_b(\eta) \\ &\quad = \sum_{(A_0,\ldots,A_n)} \sum_{(B_0,\ldots,B_n)} \pi_{b,x_0}(A_0) P^{A_0}(A_t^{b,x_0} = B_0) \cdots \pi_{b,x_n}(A_n) P^{A_n}(A_t^{b,x_n} = B_n) \times \\ &\qquad\qquad g(A_0, \ldots, A_n) f(B_0, \ldots, B_n). \end{aligned}$$

Using the fact that $\pi_{b,x_i}(A_i)P^{A_i}(A_t^{b,x_i} = B_i) = \pi_{b,x_i}(B_i)P^{B_i}(A_t^{b,x_i} = A_i)$,

$$\int_X g(\eta)\ S(t)f(\eta)\ \mathrm{d}\mu_b(\eta)$$

$$= \sum_{(A_0,\ldots,A_n)} \sum_{(B_0,\ldots,B_n)} \pi_{b,x_0}(B_0)P^{B_0}(A_t^{b,x_0} = A_0)\cdots\pi_{b,x_n}(B_n)P^{B_n}(A_t^{b,x_n} = A_n) \times$$

$$g(A_0,\ldots,A_n)f(B_0,\ldots,B_n)$$

$$= \int_X f(\eta)\ S(t)g(\eta)\ \mathrm{d}\mu_b(\eta).$$

So it follows that μ_b is reversible. ∎

Remark. Observe that there was nothing special about the fact that the initial state was taken to be δ_b in paragraph one. In fact, $\delta_\eta S(t)$ converges in distribution to μ_b for all $\eta \in X_b$. Therefore, μ_b is the unique invariant measure on X_b.

The goal is to show that the mapping from backbones to extremal reversible measures is onto. It is here that Theorem 5.1.3 comes into play. In fact, we will see that as a consequence of Theorem 5.1.3, reversible measures concentrate on configurations that are connected. Once that is established, semi-infinite configurations (infinite configurations that do not contain a backbone) need to be excluded. It turns out that semi-infinite configurations die out in a distributional sense (see Lemma 5.2.3 below). Combining the connectedness together with the fact that semi-infinite configurations die out and the fact that the state spaces X_b and $X_{b'}$ do not communicated for $b \neq b'$, will give the desired result.

Lemma 5.2.2 *If $\mu \in \mathcal{R}$, then the support of μ is contained in $X' = \{\eta \in X :$*

η *is connected* $\}$.

Proof. Consider $L_i = \{x_0, \ldots, x_i\} \subseteq \mathbb{T}^d$ be such that $\|x_j - x_{j+1}\| = 1$ for $1 \leq j < i$ and $x_k \neq x_j$ for $k \neq j$. Let $f_i(\eta) = \mathbf{1}_{\{\eta(x_0)=\eta(x_i)=1, \eta(x_j)=0, 1\leq j\leq i-1\}}$. Since $c(x_1, \eta) = 0$ whenever $f_2(\eta_{x_1}) = 1$, it follows that

$$\int_X c(x_1, \eta) f_2(\eta_{x_1}) \mathrm{d}\mu(\eta) = 0.$$

By reversibility,

$$\int_X c(x_1, \eta) f_2(\eta) \mathrm{d}\mu(\eta) = 0.$$

On the other hand,

$$\int_X c(x_1, \eta) f_2(\eta) \mathrm{d}\mu(\eta) \geq 2\beta\mu(\eta : \eta(x_0) = 1, \eta(x_1) = 0, \eta(x_2) = 1).$$

Therefore, $\mu(\eta : \eta(x_0) = 1, \eta(x_1) = 0, \eta(x_2) = 1) = 0$. In particular, all distinct connected components are at least at distance three from each other μ almost surely.

Proceeding inductively, assume that all distinct connected components are at least at distance n from each other. Thus, $f_n(\eta_{x_1}) = 0$ μ almost surely so that

$$\int_X c(x_1, \eta) f_n(\eta_{x_1}) \mathrm{d}\mu(\eta) = 0.$$

By reversibility,

$$\int_X c(x_1, \eta) f_n(\eta) \mathrm{d}\mu(\eta) = 0.$$

On the other hand,

$$\int_X c(x_1, \eta) f_n(\eta) \mathrm{d}\mu(\eta) \geq \beta\mu(\eta : \eta(x_0) = \eta(x_n) = 1, \eta(x_j) = 0, 1 \leq j \leq n-1).$$

Therefore, $\mu(\eta : \eta(x_0) = \eta(x_n) = 1, \eta(x_j) = 0, 1 \leq j \leq n-1) = 0$. In particular, all distinct connected components are at least at distance $n+1$ from each other μ almost surely.

We have shown that the μ probability that there are distinct connected components that are at a finite distance from each other is zero. In other words, $\mu(\eta : \eta \text{ is connected}) = 1$, where $\eta = \emptyset$ is considered to be connected. ∎

Lemma 5.2.3 *Let $\sigma \in X$ be a connected configuration that does not contain a backbone. Then $\delta_\sigma S(t) \to \delta_\emptyset$.*

Proof. Since $\beta \leq \beta_1$, the assertion holds for finite connected configurations σ. Fix σ infinite. Given $x \in \mathbb{T}^d$, let y_0 be the unique $y \in \sigma$ such that $\|x - y_0\| = \min\{\|x-y\| : y \in \sigma\}$. Also, let $\{y_n\}_{n\in\mathbb{N}}$ be an infinite path in σ such that $y_i \neq y_j$ for $i \neq j$ and $\|y_i - y_{i+1}\| = 1$ for all $i \in \mathbb{N}$. Finally, let $\eta_t^{n,\sigma}$ be the uniform model with initial state σ and death at y_n suppressed. By attractiveness, we can couple η_t^σ and $\eta_t^{n,\sigma}$ such that $\eta_t^\sigma \subseteq \eta_t^{n,\sigma}$ for all $t \geq 0$. Therefore,

$$P(x \in \eta_t^\sigma) \leq P(x \in \eta_t^{n,\sigma}). \tag{5.2.2}$$

Let $X_n = \{z \in \mathbb{T}^d : \|z - y_{n-1}\| \leq \|z - y_n\|\} \cup y_n$ and $A_t^n = \eta_t^{n,\sigma} \cap X_n$. Thus A_t^n is distributed as a rooted chain with $A_0^n = \sigma \cap X_n$. In particular, for $m = n + \|y_0 - x\|$, we have

$$\begin{aligned}\lim_{t\to\infty} P(x\in\eta_t^{n,\sigma}) &= \lim_{t\to\infty} P(x\in A_t^n)\\ &= \sum_{B\ni x}\frac{\beta^{|B|}}{C(\beta)}\\ &= \sum_{k_1=0}^{\infty}\cdots\sum_{k_{(d-1)m+1}=0}^{\infty}\frac{\beta^m}{C(\beta)}\beta^{k_1}\cdots\beta^{k_{(d-1)m+1}}\\ &= \frac{\beta^m}{C(\beta)}C(\beta)^{(d-1)m+1}\\ &= \left(\beta C(\beta)^{d-1}\right)^m .\end{aligned}$$

Consequently, $\beta C(\beta)^{d-1}\le 1$. Recall equation (2.5.3) that says that $C(\beta)$ and β are related via the polynomial expression

$$\beta C(\beta)^d - C(\beta) + 1 = 0$$

whenever $\beta\le\beta_1$. If $\beta C(\beta)^{d-1}=1$, this polynomial expression would reduce to $1=0$, a contradiction. Therefore, $\beta C(\beta)^{d-1}<1$ so that $\lim_{t\to\infty}P(x\in\eta_t^{n,\sigma})$ tends to zero as n tends to infinity. By equation (5.2.2),

$$\limsup_{t\to\infty} P(x\in\eta_t^{\sigma})\le\lim_{n\to\infty}\lim_{t\to\infty}P(x\in\eta_t^{n,\sigma})=0,$$

from which the result follows. ∎

Together Lemmas 5.2.2 and 5.2.3 imply the following corollary:

Corollary 5.2.4 *If $\mu\in\mathcal{R}$ and $\mu\perp\delta_{\emptyset}$, then $\mu(\eta:\exists$ a backbone $b\ni b\subseteq\eta)=1$.*

Corollary 5.2.4 together with fact that X_b and $X_{b'}$ do not communicate for $b \neq b'$ will imply that each nontrivial extremal reversible measure is given by μ_b for some backbone b. As a consequence of the fact that X_b and $X_{b'}$ do not communicate for $b \neq b'$, $\mu \in \mathcal{J}_e$ implies that a certain zero-one law holds (see Lemma 5.2.5 below). The reason why the desired result does not follow immediately from the corollary and the zero-one law is that the number of backbones is uncountable. Therefore, under the assumption that $\mu \in \mathcal{R}_e$, there is no a priori guarantee that $\mu(X_b) > 0$ for some b. This technical difficultly can be circumvented by considering the event that a particular vertex x is in the backbone. Because there are a countable number of vertices, some vertex is in the backbone with positive probability. The zero-one law then allows a single backbone to be singled out.

Lemma 5.2.5 *If $\mu \in \mathcal{R}_e$ and $Y = \cup_{b\in\mathbb{I}} X_b$ for some subset of backbones $\mathbb{I}$, then $\mu(Y) = 0$ or $\mu(Y) = 1$.*

Proof. Fix $\mu \in \mathcal{R}_e$. Suppose that there exists a subset of backbones $\mathbb{I}$ such that for $Y = \cup_{b\in\mathbb{I}} X_b$, the measure μ satisfies $0 < \mu(Y) < 1$. Let $\mu_Y(\cdot) = \mu(\cdot \mid Y)$ and $\mu_{Y^c}(\cdot) = \mu(\cdot \mid Y^c)$. Thus,

$$\mu(\cdot) = \mu_Y(\cdot)\mu(Y) + \mu_{Y^c}(\cdot)\left(1 - \mu(Y)\right).$$

Using the fact that X_b and $X_{b'}$ do not communicate for $b \neq b'$, it follows that

$S(t)\mathbf{1}_Y f = 0$ on the event that $\eta_0 \notin Y$. Therefore,

$$\int_X f\ S(t)g\ \mathrm{d}\mu_Y = \int_X \mathbf{1}_Y f\ S(t)g\ \mathrm{d}\mu_Y = \int_X \mathbf{1}_Y f\ S(t)g\ \frac{\mathrm{d}\mu}{\mu(Y)}$$
$$= \int_X g\ S(t)\mathbf{1}_Y f\ \frac{\mathrm{d}\mu}{\mu(Y)} = \int_X g\ S(t)\mathbf{1}_Y f\ \mathrm{d}\mu_Y = \int_X g\ S(t)f\ \mathrm{d}\mu_Y.$$

Hence, μ_Y is reversible. By the same argument, μ_{Y^c} is also reversible. But this contradicts the extremality of μ. Therefore, no such $\mathbb{I}$ exists. ∎

Theorem 5.2.6 $\{\mu_b : b \textit{ is a backbone}\} = \mathcal{R}_e$.

Proof. Let $\mu \in \mathcal{R}_e$ be such that $\mu \perp \delta_\emptyset$. Also, let $\mathbb{I}_x = \{b : x \in b\}$ and let $Y_x = \cup_{b\in\mathbb{I}} X_b$. By Lemma 5.2.5, either $\mu(Y_x) = 0$ or $\mu(Y_x) = 1$ for each $x \in \mathbb{T}^d$. Furthermore, by Lemma 5.2.4, $\mu(\cup_x Y_x) = 1$ so it follows that $\mu(Y_x) = 1$ for some $x \in \mathbb{T}^d$. Therefore, the set $r = \{x : \mu(\eta : x \in \eta) = 1\}$ is nonempty. In addition, the set r is connected by Lemma 5.2.2. If r is not a backbone, then r contains a leaf ℓ that has d neighbors ℓ_i, $i = 1, \dots, d$, that are in the complement of r. Since $\ell \in r$, it follows that $\mu(\cup_{i=1}^d Y_{\ell_i}) = 1$. By Lemma 5.2.5, it must be that $\mu(Y_{\ell_i}) = 1$ for at least one index i. But this contradicts the fact that $\ell_i \notin r$. Therefore, r is a backbone. Furthermore, for each $x \in \partial r$, $\mu(Y_x) = 0$ (otherwise $x \in r$). Therefore, r is the maximal backbone μ almost surely so that $\mu(X_r) = 1$. Since the uniform model restricted to the space X_r has a unique reversible measure, it follows that $\mu = \mu_r$. ∎

Note that Theorem 1.4.5 is less precise version of Theorem 5.2.6. As mentioned in the introduction, it would be interesting to determine whether or not there are any invariant measures that are not reversible. The crucial place where reversibility was used in this section was to prove the connectedness property. The remainder of the proof relies on connectedness, not reversibility. So the main issue is to determine whether or not there are invariant measures that are supported on collections of disconnected configurations.

Bibliography

[1] E. Andjel. The contact process in high dimensions. *Ann. Probab.*, 16:1174–1183, 1988.

[2] M. Barlow, R. Pemantle, and E. Perkins. Diffusion-limited aggregation on a tree. *Probab. Theory Relat. Fields*, 107:1–60, 1997.

[3] D. Barsky and C. Wu. Critical exponents for the contact process under the triangle condition. *to appear in J. Statist. Phys.*, 1998.

[4] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.*, 18:984–1009, 1991.

[5] D. Chen. On the survival probability of generalized nearest particle systems. *Stoch. Proc. Appl.*, 30:209–223, 1988.

[6] D. Chen. Finite nearest particle systems on a tree. *Acta Math. Scienta*, 14:348–353, 1994.

[7] L. Devroye. A note on the height of binary search trees. *Journal ACM*, 33:489–498, 1986.

[8] P. Doyle and J. Snell. *Random walk and electric networks.* Mathematical Association of America, Washinton D.C., 1984.

[9] R. Durrett. On the growth of one dimensional contact processes. *Ann. Probab.*, 8:890–907, 1982.

[10] R. Durrett and D. Griffeath. Contact processes in several dimensions. *Z. Wahrsch. Verw. Gebiete*, 59:535–552, 1982.

[11] R. Durrett and R. Schinzai. Intermediate phase for the contact process on a tree. *Ann. Probab.*, 23:668–673, 1995.

[12] P. Gill, W. Murray, and M. Wright. *Numerical Linear Algebra and Optimization.* Addison-Wesley Publishing Company, 1991.

[13] D. Griffeath and T. Liggett. Critical phenomenon for Spitzer's reversible nearest-particle system. *Ann. Probab.*, 10:881–895, 1982.

[14] T. Harris. Contact interactions on a lattice. *Ann. Probab.*, 2:652–680, 1974.

[15] P. Lalley and T. Sellke. Limit set of a weakly supercritical contact process on a homogeneous tree. *to appear in Ann. Probab.*, 1998.

[16] T. Liggett. Attractive nearest particle systems. *Ann. Probab.*, 11:16–33, 1983.

[17] T. Liggett. *Interacting Particle Systems.* Springer-Verlag, 1985.

[18] T. Liggett. Reversible growth models on symmetric sets. *Proceedings of the Taniguchi Symposium*, pages 275–301, 1985.

[19] T. Liggett. Reversible growth models on $\mathbb{Z}^d$: some examples. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, pages 213–227. Springer-Verlag, 1987.

[20] T. Liggett. Branching random walks and contact processes on homogeneous trees. *Probab. Theory Relat. Fields*, 106:495–519, 1996.

[21] T. Liggett. Multiple transition points for the contact process on the binary tree. *Ann. Probab.*, 24:1675–1710, 1996.

[22] T. Liggett. *Contact, Voter, and Exclusion Processes.* preliminary version, 2000.

[23] T. Lyons. A simple criterion for transience of a reversible Markov chain. *Ann. Probab.*, 11:393–402, 1983.

[24] T. Mountford. A complete convergence theorem for attractive reversible nearest particle systems. *Canad. J. Math.*, 49:321–337, 1997.

[25] R. Pemantle. The contact process on trees. *Ann. Probab.*, 20:2089–2116, 1992.

[26] B. Pittel. On growing random binary trees. *J. Math. Anal. Appl.*, 103:461–480, 1985.

[27] M. Salzano and R. Schonmann. The second lowest extremal invariant measure of the contact process. *Ann. Probab.*, 25:1846–1871, 1997.

[28] M. Salzano and R. Schonmann. A new proof that for the contact process on homogeneous trees local survival implies complete convergence. *to appear in Ann. Probab.*, 1998.

[29] M. Salzano and R. Schonmann. The second lowest extremal invariant measure of the contact process II. *to appear in Ann. Probab.*, 1998.

[30] R. Schonmann. A new proof of the complete convergence theorem for contact processes in several dimensions with large infection parameter. *Ann. Probab.*, 15:382–387, 1987.

[31] R. Schonmann. The triangle condition for contact processes on homogeneous trees. *preprint*, 1998.

[32] F. Spitzer. Stochastic time evolution of one-dimensional infinite-particle systems. *Bull. Amer. Math. Soc.*, 83:880–890, 1977.

[33] A. Stacey. The existence of an intermediate phase for the contact process on trees. *Ann. Probab.*, 24:1711–1726, 1996.

[34] C. Wu. The contact process on a tree: Behavior near the first transition. *Stoch. Proc. Appl.*, 57:99–112, 1995.

[35] Y. Zhang. The complete convergence theorem of the contact process on trees. *Ann. Probab.*, 24:1408–1443, 1996.