The Segregated Λ-coalescent

Nic Freeman

St Anne's College
Oxford University

Thesis submitted for the degree of Doctor of Philosophy in Mathematics, 2nd August 2012.

Acknowledgement

I would like to thank my supervisors Profs. Alison Etheridge and Ben Hambly for their advice and support throughout all stages of my D.Phil. I would also like to thank Prof. Terry Lyons for his help with the C++ code that began as a toy project and later led me to investigate the Segregated Λ-coalescent.

Lastly, to the many people who offered me invaluable help and advice during my time at Oxford; I am grateful to you all.

Abstract

We study a natural generalization of the Λ-coalescent to a spatial continuum. We introduce the process, which is known as the Segregated Λ-coalescent, via its connections to the (non-spatial) Λ-coalescent and the Spatial Λ-Fleming-Viot process.

The main new results contained in this thesis are as follows. The Segregated Λ-coalescent has a non-trivial construction which we present here in terms of stochastic flows. We describe the qualitative behaviour of the Segregated Λ-coalescent and compare it to the behaviour of the Λ-coalescent, showing in particular that the Segregated Λ-coalescent has an extra phase transition which is directly related to the introduction of space. We finish with some results concerning the rate at which the Segregated Λ-coalescent comes down from infinity.

Contents

Chapter 1

Introduction

Coalescent processes are stochastic models in which a collection of particles start out separated and then coalesce over time. The modern theory begins with the coalescent of Kingman (1982), which has been studied intensively and extended to give a rich family of coalescent models.

The theory of coalescent processes is closely linked to mathematical biology, in particular to population modelling. A population model is a set of rules which describe the reproduction of individuals within some idealized population. It is usual to consider models in which each individual has only one parent. Thus, the model endows the set of individuals with a tree structure, in which individuals are nodes and a single edge connects together each parent-child pair. This tree structure is known as the genealogy of the population; it formalizes the notion of a family tree[1].

Frequently, situations occur in which a coalescent process may be used to describe the genealogy, viewed backwards in time, of individuals whose reproduction is given by some population model. Kingman's coalescent is a universal object in this respect and has been shown to arise as the (rescaled limit of the) genealogies of a number of classical population models. A canonical example is the relationship between Kingman's coalescent and the Wright-Fisher model, see e.g. Etheridge (2011).

The development of modern coalescent theory arose out of a desire to incorporate more realistic features into Kingman's coalescent. In particular, the Λ-coalescent was introduced (independently, but in the same spirit) by Donnelly and Kurtz (1999), Pitman (1999) and Sagitov (1999). An excellent introduction to the theory of Λ-coalescents can be found in

[1] In fact, some of the first population models were developed by the Victorians to estimate the probability of aristocratic surnames becoming extinct.

Berestycki (2009).

We will discuss the Λ-coalescent in Chapter 2, but for now let us give a brief heuristic outline of the process. At time 0 the Λ-coalescent begins with a countable infinity of particles, with each particle representing an individual from the population. It is usual to label these particles with elements of $\mathbb{N}$. Then, at a countable set of random times during $(0, \infty)$ a subset of the currently present particles are selected and these particles come together to form a coalesced block of particles. This coalesced block is thought of as a single new particle and may subsequently be coalesced into even larger blocks of particles.

For the moment we do not need to specify the mechanism that determines precisely which particles participate in which coagulation events. However, we should mention that the resulting coalesced blocks of particles are either singletons or infinite sets. From the perspective of a single particle, it begins alone as a singleton and remains so until it is affected by a coalescence event, at which point it is subsumed into an ever growing block that contains infinitely many other particles. In a sense that will be made rigorous later on, each of these infinite blocks contains a non-zero proportion of the initial particles.

It is possible for coalescence to occur sufficiently fast that, with probability one, at all times $t > 0$ only a finite number of (coalesced blocks of) particles remain present. When this occurs we say that the coalescent comes down from infinity. In such cases, denoting the total number of blocks at time $t \geq 0$ by $|\Pi_t|$, the asymptotic rate at which $|\Pi_t| \uparrow \infty$ as $t \downarrow 0$ is known as the rate of coming down from infinity. This rate has been the subject of much interest over the past decade and will be discussed in Chapter 2.

In population genetics the term 'spatial' usually refers to the geographical space, which is the space in which individuals of the population are thought to live, feed, move around, etc. A non-spatial model is one in which the geographical space has no influence on the dynamics of the model and is effectively disregarded.

The Λ-coalescent is a non-spatial model, which means that the random forces which cause particles to coalesce are felt evenly across the entire system. Informally, each pair (triplet, etc) of particles is as likely to coalesce as any other pair (triplet, etc). In Chapter 2 we will refer to this property of the Λ-coalescent as *exchangeability*. Exchangeability is not a realistic assumption with regards to modelling genealogies, but from a mathematical point of view it is an extremely useful feature which promotes tractability of the model.

1.1 Our results

The focus of this thesis is the Segregated Λ-coalescent, a new model which offers a method of extending the Λ-coalescent to spatial continua. We will sometimes refer to the Segregated Λ-coalescent as 'our model'.

The Segregated Λ-coalescent preserves the spirit of the Λ-coalescent but drops the exchangeability. The result is a coalescent process in which each particle lives at some point in a continuum of geographical space. Particles that start close together will typically coalesce faster than those that start far apart.

We introduce the Segregated Λ-coalescent through its connections to the Λ-coalescent and the recently introduced Spatial Λ-Fleming-Viot (SΛFV) process. The Λ-coalescent and SΛFV process are discussed in Chapters 2 and 3 respectively.

Our model is formulated as a stochastic flow and has a non-trivial construction which is given in Chapter 4. Using the flow we show that the Segregated Λ-coalescent is dual to a population model which is itself a close relative of the SΛFV process.

Like the Λ-coalescent, the blocks of the Segregated Λ-coalescent are either singletons (collectively known as dust) or infinite blocks that contain a non-trivial proportion of the population. We say that a coalescent experiences a phase transition when a small variation in the parameters of the model causes a significant change in the qualitative behaviour of the dust and/or non-singleton blocks.

Our model exhibits one especially notable feature when compared to the SΛFV process: in the Segregated Λ-coalescent particles can coalesce much faster than in the (dual of the) SΛFV process. The material consequence of this is that the Segregated Λ-coalescent is able to come down from infinity, whereas the dual of the SΛFV process is not.

The ability to come down from infinity is not achieved easily. Our models definition is tailored to allow very fast coalescence and, since particles must move through space when they coalesce, we define our particle motions in such a way that a large number of small jumps cannot make a single particle travel across a great distance in space. A key element of achieving this is a tree-like structure for the geographical space. This structure provides a means of controlling how far particles move when they jump.

We give a complete classification of the phases of the Segregated Λ-coalescent in Chapter

5, showing that many of the distinctive features of the Λ-coalescent are preserved through the introduction of space and are expressed by the Segregated Λ-coalescent. However, the introduction of space also enriches the behaviour of the coalescent through the appearance of a 'critical phase' in which our model comes down from infinity gradually, over a deterministic bounded time interval.

Our model has a close connection, which is described in Chapter 6, to Galton-Watson processes in varying environments. The results stated in Chapter 5 are proved in Chapter 7, making heavy use of this connection. Then, in Chapter 8 we obtain more detailed information concerning the behaviour of our model in its critical phase.

In Chapter 9 we study the asymptotic rate at which the Segregated Λ-coalescent comes down from infinity. We are not able to give fully general results and we focus on a (large) class of cases in which the rate of coalescence is significantly faster than is needed to make our model come down from infinity. In such cases we are able to give formulae for the asymptotic rate at which the Segregated Λ-coalescent comes down from infinity.

1.2 The Segregated Λ-coalescent

In this section we give an informal description of the Segregated Λ-coalescent, focused on the special case in which the geographical space is taken to the be Cantor set. This case may be regarded as a prototype case of our model. We write $A \uplus B$ to denote the disjoint union[2] of A and B.

Let $\mathcal{S}$ be some integer greater than or equal to 2 and let $S = \{1, 2, \dots \mathcal{S}\}$. Recall that the $\mathcal{S}$-part Cantor set K is the unique non-empty compact subset of $[0, 1]$ which satisfies $K = \biguplus_{i=1}^{\mathcal{S}} F_i(K)$ where $F_i(x) = \frac{1}{2\mathcal{S}-1}(2i - 2 + x)$. If $w = w_1 \dots w_n$, where $w_j \in S$, then we call the set

$$K_w = F_{w_1} \circ \dots \circ F_{w_n}(K)$$

an n-complex of K, or, when n is not specified, a complex of K. The 1-complexes $K_i = F_i(K)$ consist of $\mathcal{S}$ copies of K, each shrunk by a factor of $\frac{1}{2\mathcal{S}-1}$, evenly spaced across $[0, 1]$. If $K_w \subseteq K_{w'}$ then we say $K_{w'}$ is a subcomplex of K_w.

Let W_n be the set of words of length $n \in \mathbb{N}$ with letters in S (formally, $W_n = S^n$ but we will always write $w \in W_n$ as $w = w_1 w_2 \dots w_n$ where $w_j \in S$). If $w = w_1 \dots w_n \in W_n$ then

[2] That is, $A \uplus B = A \cup B$ with the implication that $A \cap B = \emptyset$.

the length of the word w is $|w| = n$. Let $W_0 = \{\emptyset\}$ and $|\emptyset| = 0$, which says that $\emptyset$ is the empty word of length 0. Set

$$W_* = \bigcup_{n=0}^{\infty} W_n.$$

As usual, K is the disjoint union of its n-complexes:

$$K = \biguplus_{w \in W_n} K_w. \tag{1.2.1}$$

The most important parameter in the model is a sequence $(r_n)_{n \geq 0}$. This sequence will control how fast the particles within our model coalesce. To avoid triviality we require that $r_n > 0$ for some n.

Let $\mathcal{U}$ denote the uniform probability measure on K, so that $\mathcal{U}(K_w) = \mathcal{S}^{-|w|}$. Let $\mathcal{U}_w$ denote the conditional measure of $\mathcal{U}$ on K_w, defined by $\mathcal{U}_w(A) = \frac{\mathcal{U}(A \cap K_w)}{\mathcal{U}(K_w)}$. Let $\mathcal{R}$ denote the measure on the countable set W_* which is defined via the point-masses $\mathcal{R}(w) = r_{|w|}$. Let M be a Poisson point process which has points $(t, w, p) \in (-\infty, \infty) \times W_* \times K$ and rate

$$dt \otimes \mathcal{R}(dw)\mathcal{U}_w(dp). \tag{1.2.2}$$

Definition 1.2.1 (Segregated Λ-coalescent, informal) *We define a stochastic flow $(X_{s,t})_{-\infty < s < t < \infty}$ on K as follows.*

- *Whenever $(t, w, p) \in M$, any particles in the flow which are in K_w at time $t-$ jump to p at time t.*
- *The position of each particle is constant in between its jumps.*

When $(t, w, p) \in M$, we say a reproduction event has occurred in K_w at time t with parent p. Often we will shorten this to 'level n reproduction event', where $|w| = n$. Figure 1.1 gives a graphical demonstration of Definition 1.2.1 in the case of $X_{0,t}$ when $\mathcal{S} = 2$.

Definition 1.2.1 makes rigorous sense if, and only if, $\sum_0^\infty \mathcal{S}^n r_n < \infty$. This condition characterises the case where only finitely many reproduction events occur during any bounded interval $[s, t]$. In Chapter 4 we will use a novel method to make sense of Definition 1.2.1 for *any* sequence $(r_n)_{n \in \mathbb{N}_0} \subseteq [0, \infty)$.

Remark 1.2.2 *Throughout this thesis we will encounter stochastic processes which are described in terms of a Poisson point process. See Appendix A for a brief discussion of such informal descriptions.*

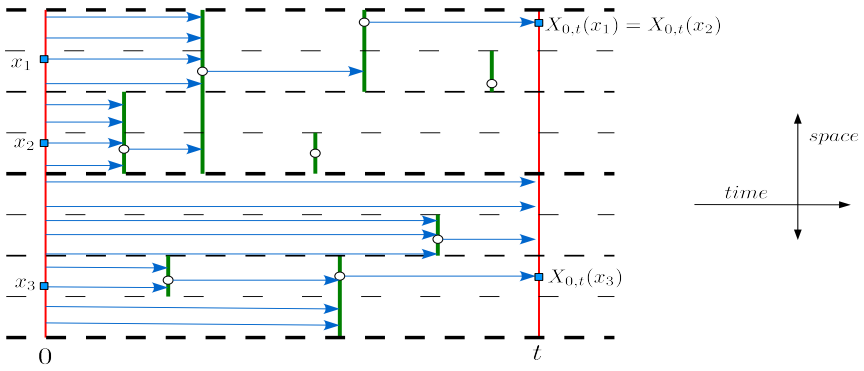


Figure 1.1: The complexes of the geographical space with $\mathcal{S} = 2$ are shown down to level 3, with dotted lines. Note that the space between neighbouring complexes has been removed. In this realisation there are no reproduction events occurring in complexes of level 4 and above. The movement of some sample particles over $[0, t]$ is indicated by arrows, with the start/end-points from x_1, x_2, x_3 marked with square dots. Reproduction events are shown as thick vertical lines, with parents as circular dots.

The Segregated Λ-coalescent is not restricted to totally disconnected geographical spaces like the Cantor set. We are able to deal with situations where the complexes touch (that is, where $\overline{K_w} \cap \overline{K_{w'}} \neq \emptyset$ and $K_w \cap K_{w'} = \emptyset$).

Definition 1.2.1 suggests that our model has a close connection to Galton-Watson processes in varying environments (GWVEs). Recall that a GWVE is a classical Galton-Watson process with the modification that the offspring distribution of an individual may depend on its generation number. The connection is as follows: for $w \in W_*$ define

$$\mathcal{E}_w = \inf\{t > 0\,;\, \exists u \in W_*, p \in K \text{ such that } K_w \subseteq K_u \text{ and } (t, u, p) \in M\}. \tag{1.2.3}$$

We refer to $\mathcal{E}_w$ as the exponential clock associated to K_w. For each $t > 0$ and $n \geq 0$ define

$$\mathscr{B}_n^t = \big\{w \in W_n\,;\ \text{for all } j = 0, 1, \ldots, n,\ \mathcal{E}_{w_1 \ldots w_n} > t\big\}. \tag{1.2.4}$$

Set $B_n^t = |\mathscr{B}_n^t|$. Then, for each fixed $t > 0$, $(B_n^t)_{n \geq 1}$ is a GWVE. To see this, note first that $e^{-r_n t}$ is the probability that K_w, where $|w| = n$, does not see its clock ring during $(0, t]$. If $w \in \mathscr{B}_n^t$ and $|w| = n$, then the (conditional) probability that $wi \in \mathscr{B}_{n+1}^t$ is just $e^{-r_{n+1} t}$. The clocks corresponding to K_{wi} and K_{wj} are independent if $i \neq j$, thus the offspring distribution of $w \in \mathscr{B}_n^t$ is binomial with $\mathcal{S}$ trials and success probability $e^{-r_{n+1} t}$.

The connections between the Segregated Λ-coalescent and GWVEs are described in detail in Chapter 6. They form the basis of our analysis of the models behaviour.

1.3 Some notation

All the spaces we consider will be metric spaces and, unless explicitly stated otherwise (which means, except in Appendix C), we use the topology induced from the metric. We equip all metric spaces with the Borel σ-field induced from their topology. We write $\mu(dx) \otimes \nu(dy)$ for the product measure of μ and ν, with the usual convention that $dx, dy,$ etc refers to the Lebesgue measure on $\mathbb{R}^d$. The measure corresponding to a single point-mass at x is written δ_x.

We write integrals as $\int f(x)\, dx$, except when we deal with Markov generators whose formulae require a significant number of repeated integrals. In these cases we will tend to use the operator notation $\int dx \int dy \int dz\, (\ldots)$ instead of $\int\int\int (\ldots) dx\, dy\, dz$.

We write $|A|$ for the number of elements of the set A, as well as writing $|a|$ to denote the Euclidean norm on $\mathbb{R}^d$. When A is a suitable space of functions we write $||a||_\infty$ and $||a||_p$ for the appropriate supremum and p-norms respectively of $a \in A$. For suitable spaces A, we write $C(A)$ for the space of continuous real valued functions on A, equipped with $||\cdot||_\infty$ norm and corresponding metric.

When we say '... events occur at rate λ', we mean that such events occur at random times with inter-arrival times given by independent exponential random variables with mean $\frac{1}{\lambda}$.

Superscripts may denote powers or indexes, depending upon the context. We will always take care in our definitions to specify which objects have which indexes. For $y \in \mathbb{R}$ we write $x \uparrow y$ and $x \downarrow y$ for the left and right limits respectively, as $x \to y$.

We write $\min(a, b) = a \wedge b$ and $\max(a, b) = a \vee b$. We use the conventions that $\sum_{j=n}^{m} \ldots = 0$ and $\prod_{j=m}^{n} \ldots = 1$ if $m < n$.

The symbol M will refer to different Poisson point processes over the course of this thesis. Our convention is that M always denotes the Poisson point process corresponding to the model which is currently in focus. Thus M corresponds to the Λ-coalescent and SΛFV process in Chapters 2 and 3 respectively, whereas in all other chapters M corresponds to the Segregated Λ-coalescent.

Remark 1.3.1 *A diagram of the dependencies between the sections and an index of the main pieces of our notation can both be found after the appendices.*

Chapter 2

The Λ-coalescent

2.1 Definition

Let $\mathcal{P}_n$ denote the set of partitions of $\{1, 2, \ldots, n\}$. Let ι_n denote the natural restriction map $\iota_n : \mathcal{P}_{\mathbb{N}} \to \mathcal{P}_n$, which is defined by simply removing all elements $m \in \mathbb{N} \setminus \{1, \ldots, n\}$ from a partition of $\mathbb{N}$. For example, $\iota_3(\{(1,2,6),(3,5),(4),\ldots\}) = \{(1,2),(3)\}$. Let $1^{\mathbb{N}}$ denote the partition of $\mathbb{N}$ into singletons.

If π is a partition of $\mathbb{N}$, then each element of π is known as a *block*. We write $n \overset{\pi}{\sim} m$ to mean that n and m are in the same block of π. If $\pi = \{b_1, b_2, \ldots b_l\}$ (resp. $\pi = \{b_i\,;\, i \in \mathbb{N}\}$) and $I \subseteq \{1, \ldots, l\}$ (resp $I \subseteq \mathbb{N}$) then the partition obtained from π by *merging* $\{b_i\,;\, i \in I\}$ is given by $\{b_i\,;\, i \notin I\} \cup \{\cup_{i \in I} b_i\}$. We use the terms merging and *coagulating* interchangeably.

Definition 2.1.1 (The Λ-coalescent) *Let Λ be a finite measure on $[0,1]$. The Λ-coalescent is the (unique) $\mathcal{P}_{\mathbb{N}}$-valued Markov process $(\Pi_t)_{t \geq 0}$ such that, for all $n \in \mathbb{N}$, $\Pi_t^{(n)} = \iota_n(\Pi_t)$ is a $\mathcal{P}_n$-valued Markov chain with initial state $1^{\mathbb{N}}$ and the following dynamics: Whenever $\Pi_t^{(n)}$ is a partition consisting of i blocks, the rate at which any k-tuple of blocks merges is*

$$\lambda_{i,k} = \int_0^1 x^{k-2}(1-x)^{i-k}\Lambda(dx), \tag{2.1.1}$$

independently of all other k-tuples.

The formulation of Definition 2.1.1 is due to Pitman (1999). If $\Lambda(\{0\}) = 0$, then the formula (2.1.1) is more intuitively written as $\lambda_{b,k} = \int_0^\infty x^k(1-x)^{b-k}x^{-2}\Lambda(dx)$. The term

$$\nu(dx) = x^{-2}\Lambda(dx) \tag{2.1.2}$$

corresponds to a measure controlling the rate at which a proportion $x \in (0,1]$ of the blocks currently present merge to form a new block. The remaining 'binomial' term $x^k(1-x)^{b-k}$

says that, of the first b blocks, each block chooses independently whether to become part of the new block or to remain alone (with probabilities x and $1-x$ respectively).

If $\Lambda(\{1\}) > 0$ then corresponding events occur, at rate $\Lambda(\{1\})$, which coagulate the whole population into a single block. From a theoretical point of view, this adds no extra complexity and serves only to obfuscate the behaviour which we will later describe. Results concerning Λ-coalescents for which $\Lambda(\{1\}) = 0$ are easily extended to general Λ by simply superimposing the extra coagulation events. With this in mind:

Remark 2.1.2 *For the remainder of Chapter 2 we consider only Λ for which $\Lambda(\{1\}) = 0$.*

For us, it is most important is to understand the character of the process; the Λ-coalescent is a $\mathcal{P}$-valued Markov process in which, at random points in time, a selection of the blocks that are currently present merge into a single block.

Kingman's coalescent corresponds to the special case where Λ is a point-mass at 0. In this case, as is readily seen from (2.1.1), each merger involves only a pair of blocks. In general, the Kingman component of the Λ-coalescent is the atom of Λ at 0, which causes a Kingman coalescent of rate $\Lambda(0)$ to be superimposed upon the other coagulation events.

Remark 2.1.3 *In this thesis we will not consider Ξ-coalescents, which are $\mathcal{P}_{\mathbb{N}}$-valued coalescent processes in which more than one new block may be created at the same instant of time. The family of Ξ-coalescents, which further generalize Kingman's coalescent, were introduced by Schweinsberg (2000b) and Möhle and Sagitov (2001).*

It is usual to equip the set of blocks of the Λ-coalescent with a genealogy, defined as follows. We say a block A, created in a merger at time t from blocks $A = \bigcup_{i\in I} A_i$, is a parent with children $\{A_i\,;\, i \in I\}$. Thus each realization of $t \mapsto \Pi_t$ gives rise to a natural tree structure on the set $\cup_t \Pi_t$ and on the restrictions $\cup_t \Pi_t^{(n)}$.

2.2 Coming down from infinity

There is one especially prominent feature in the behaviour of the Λ-coalescent, which we now discuss.

Definition 2.2.1 *The Λ-coalescent is said to come down from infinity if*

$$\mathbb{P}\left[\forall t > 0, |\Pi_t| < \infty\right] = 1.$$

When the Λ-coalescent does come down from infinity considerable effort has been devoted to establishing the asymptotic growth of $|\Pi_t|$ as $t \downarrow 0$. We refer to this loosely as the rate of CDI as $t \downarrow 0$.

For the special case of Kingman's coalescent a straightforward reversal of time shows that $\frac{|\Pi_t|}{2t^{-1}} \to 1$ almost surely and in L^p as $t \downarrow 0$ (for example, as in Section 2.1.2 of Berestycki 2009). For general Λ-coalescents, a direct reversal of time is not possible but there is an elegant duality with a measure valued branching process known as the Λ-Fleming-Viot process. We will describe this duality in Section 2.5.

A condition which determines, in general, whether or not the Λ-coalescent comes down from infinity was first given by Schweinsberg (2000a). We will record a precise statement of Schweinsberg's result later on, as Theorem 2.3.5.

The first example of a (non-Kingman) Λ-coalescent appeared in Bolthausen and Sznitman (1998) and corresponds to the case where Λ is the uniform measure on $[0, 1]$. Bertoin and Le Gall (2000) discovered a correspondence between the genealogy of the Bolthausen-Sznitman coalescent and the genealogy of the continuous state branching process (CSBP) considered in Neveu (1992). This correspondence was extended to more general CSBPs and β-coalescents by Birkner et al. (2005). The β-coalescents are the subclass of Λ-coalescents in which Λ has the $\beta(2-\alpha, \alpha)$ distribution, for $\alpha \in (0, 2)$. The $\beta(1, 1)$ distribution is uniform on $[0, 1]$, so the case $\alpha = 1$ is the Bolthausen-Sznitman coalescent.

Bertoin and Le Gall (2006) gave a condition (equivalent to that of Schweinsberg 2000a) for the Λ-coalescent to come down from infinity, in terms of the connection to CSBPs. Their condition is well known in the Lévy process literature as Grey's criterion, which is a criterion for extinction of a CSBP. Bertoin and Le Gall (2006) also studied the rate of CDI in the β-coalescent case (and perturbations thereof) in terms of convergence in probability. The convergence in probability was generalized, also using CSBP based methods, by Berestycki et al. (2008), who established the almost sure asymptotic behaviour of $|\Pi_t|$ for β-coalescents. It turns out that β-coalescents come down from infinity if and only if $\alpha \in (1, 2)$, in which case $\frac{|\Pi_t|}{t^{1/(1-\alpha)}}$ converges almost surely to a deterministic constant.

The effort to determine the rate of CDI culminated with Berestycki et al. (2010), who exhibit a function $v(t)$ such that, for a general Λ-coalescent, $\frac{|\Pi_t|}{v(t)} \to 1$ both almost surely and in L^p as $t \downarrow 0$. Berestycki et al. use a martingale method, rather than CSBPs, but the

function $v(t)$ bears a close resemblance to the formulation of the coming down from infinity condition in Bertoin and Le Gall (2006). In fact CSBPs were crucial to determining the form of the function ν and the connection is explained in Berestycki et al. (2012a).

One important property of Kingman's coalescent is that it is possible to obtain what is known, following Ewens (1972), as the Ewens sampling formula. This formula explicitly describes the distribution of the blocks (of individual with common genes) which would be obtained from sampling $n \in \mathbb{N}$ individuals (at some time $t > 0$) from a population whose genealogy follows Kingman's coalescent and a natural mutation model. The Ewens sampling formula is an important tool in mathematical biology, since it provides a method for fitting Kingman's coalescent to data.

It is not known if a generalization of the Ewens sampling formula exists in closed form for general Λ-coalescents, but the rate of CDI has been used to establish asymptotic equivalents (as the sample size $n \to \infty$) by Berestycki et al. (2012b).

Remark 2.2.2 *It is clear from the examples given above that branching processes are an important tool for studying Λ-coalescents. In later sections we will see that branching processes are also an important tool for studying the Segregated Λ-coalescent.*

2.3 Phases of the Λ-coalescent

If $f : \mathbb{N} \to \mathbb{N}$ is a bijection, and $\{b_1, b_2, \ldots\} = \pi \in \mathcal{P}_n$ (where (b_i) is possibly a finite sequence), then we define $f(\pi) = \{f(b)\,;\, b \in \pi\}$, where $f(b) = \{f(k)\,;\, k \in b\}$.

Definition 2.3.1 *A random partition π taking values in $\mathcal{P}_\mathbb{N}$ is said to be exchangeable if, for any permutation σ of $\mathbb{N}$, $\sigma(\pi)$ has the same distribution as π.*

Pitman (1999) showed that, for any $t > 0$, Π_t is an exchangeable partition. The following lemma collects together some results, due mostly to Kingman, which can be found in Section 2.2 of Bertoin (2006).

Lemma 2.3.2 *Let π be an exchangeable partition of $\mathbb{N}$. Then, almost surely, for every block b of π, the limit*

$$\text{freq}(b) = \lim_{k\to\infty} \frac{|b \cap \{1, \ldots, k\}|}{k}$$

exists. The quantity $\mathrm{freq}(b)$ *is known as the* asymptotic frequency *of* b*. It holds that* $\sum_{b\in\pi}\mathrm{freq}(b)\leq 1$*, and also that*

$$\lim_{k\to\infty}\frac{|\{n=1,\ldots,k\,;\,\{n\}\in\pi\}|}{k}\overset{a.s.}{=}1-\sum_{b\in\pi}\mathrm{freq}(b)$$

Further, almost surely, if $\mathrm{freq}(b)=0$ *then* b *is a singleton. Almost surely,* π *has no singletons if and only if* $\sum_{b\in\pi}\mathrm{freq}(b)=1$*.*

Remark 2.3.3 *For a general subset* $b\subseteq\mathbb{N}$*, there is no reason for the limit* $\mathrm{freq}(b)$ *to exist.*

If $\mathrm{freq}(b)>0$ then the elements of b comprise a non-zero proportion $\mathrm{freq}(b)$ of $\mathbb{N}$ and we refer to b as an *non-singleton block* of π. Each singleton $\{n\}$ comprises only a null proportion of $\mathbb{N}$, in that $\mathrm{freq}(\{n\})=0$. The set of singletons of π is called the *dust* of π,and the elements of this set comprise a proportion $\mathscr{D}(\pi)\in[0,1]$ of $\mathbb{N}$, given by

$$\mathscr{D}(\pi)=1-\sum_{b\in\pi}\mathrm{freq}(b).$$

Pitman (1999) proved the following result, which establishes a dichotomy in the behaviour of the dust. Let

$$\mu^n=\int_0^1 x^n\Lambda(dx). \tag{2.3.1}$$

Recall our convention (from Remark 2.1.2) that we consider only Λ-coalescents for which $\Lambda(\{1\})=0$.

Theorem 2.3.4 (Pitman 1999) *If* $\mu^{-1}<\infty$*, then* $\mathbb{P}\left[\forall t>0,\mathscr{D}(\Pi_t)>0\right]=1$*, whereas if* $\mu^{-1}=\infty$ *then* $\mathbb{P}\left[\forall t>0,\mathscr{D}(\Pi_t)=0\right]=1$*.*

Let

$$N_t^s=|\{b\in\Pi_t\,;\,\mathrm{freq}(b)=0\}|, \tag{2.3.2}$$

which, in words, is the number of singletons of Π_t. By Lemma 2.3.2 (and the fact that $N_u^s\leq N_v^s$ when $u\geq v$), Theorem 2.3.4 implies (almost surely) that $N_t^s=\infty$ or $N_t^s=0$, and that $\mathscr{D}(\Pi_t)>0$ if and only if $N_t^s=\infty$.

A second dichotomy, this time in the behaviour of the total number of blocks of Π_t, was proved by Schweinsberg (2000a). Let

$$\mu^*=\sum_{i=2}^{\infty}\left(\sum_{k=2}^{i}(k-1)\binom{i}{k}\lambda_{i,k}\right)^{-1}.$$

Theorem 2.3.5 (Schweinsberg 2000a) *If $\mu^* < \infty$ then $\mathbb{P}[\forall t > 0, |\Pi_t| < \infty] = 1$, whereas if $\mu^* = \infty$ then $\mathbb{P}[\forall t > 0, |\Pi_t| = \infty] = 1$.*

It can be shown that $\mu^{-1} < \infty$ implies that $\mu^* = \infty$; in other words, Λ-coalescents with a non-empty dust component have infinitely many blocks for all time. In fact, we already knew this from Lemma 2.3.2 since non-empty dust implies $\mathscr{D}(\Pi_t) > 0$, which in turn implies $N_t^s = \infty$.

Remark 2.3.6 *Theorem 2.3.5 shows that the Λ-coalescent comes down from infinity if and only if $\mu^* < \infty$.*

A third (and in some sense, final) dichotomy occurs for Λ-coalescents. Let

$$N_t^a = |\{b \in \Pi_t \,;\, \text{freq}(b) > 0\}|, \tag{2.3.3}$$

which is the number of non-singleton blocks of Π_t. Note that

$$N_t^a + N_t^s = |\Pi_t|.$$

If $\mu^{-1} = \infty$ then Π_t has no singletons, so Theorem 2.3.5 tells us that $N_t^a = \infty$ when $\mu^* = \infty$ and $N_t^a < \infty$ when $\mu^* < \infty$. The behaviour of N_t^a, in the case $\mu^{-1} < \infty$, is described by the following theorem.

Theorem 2.3.7 *Suppose that $\mu^{-1} < \infty$. If $\mu^{-2} = \infty$ then $\mathbb{P}[\forall t > 0, N_t^a = \infty] = 1$, whereas if $\mu^{-2} < \infty$ then $\mathbb{P}[\forall t > 0, N_t^a < \infty] = 1$.*

Proving results about the behaviour of the Λ-coalescent is not the focus of this thesis and we give the proof of Theorem 2.3.7 in Appendix B. With Theorems 2.3.4-2.3.7 in hand, the qualitative behaviour of N^a and N^s can be completely classified, as in Figure 2.1. This divides the behaviour of the Λ-coalescent into four distinct phases.

Example 2.3.8 (β-coalescents) *Recall that the Λ-coalescent where $\Lambda(dx)$ has the $\beta(2-\alpha, \alpha)$ distribution, for $\alpha \in (0, 2)$, is known as the β-coalescent with parameter α. The case $\alpha = 1$ corresponds to the coalescent of Bolthausen and Sznitman (1998) and the case $\alpha = 2$ is by convention (or as the natural limiting case) Kingman's coalescent, where Λ is a point mass at 0.*

Label the different phases of the Λ-coalescent as $\mathscr{A}$ to $\mathscr{D}$, as in Figure 2.1. Some easy estimates show that

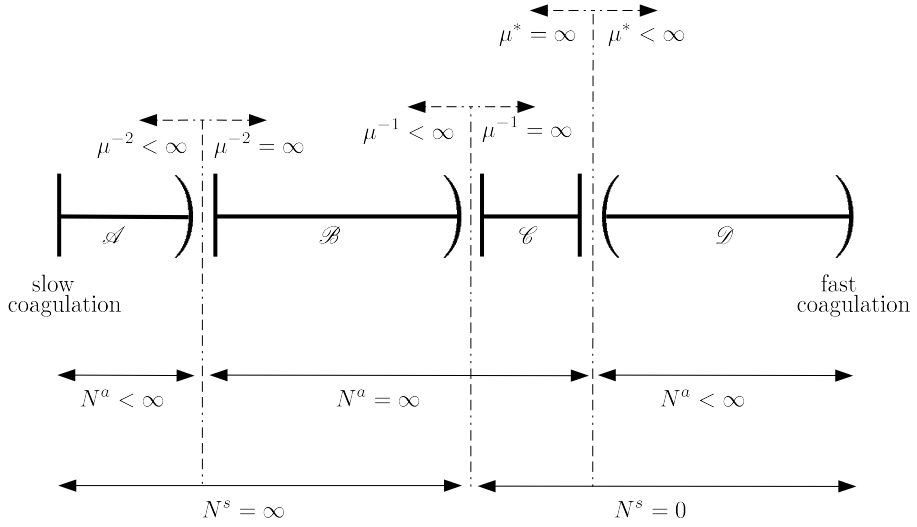


Figure 2.1: The qualitative changes in the behaviour of the Λ-coalescent are shown. The behaviour of the Λ-coalescent is determined solely by the behaviour of the measure Λ at 0+, that is, by the rate of small coagulation events. Moving from left to right, across the picture, the rate of these coagulation events increases. The resulting four behaviours are labelled $\mathscr{A}$ to $\mathscr{D}$.

- *For $\alpha \in (0, 1)$, phase $\mathscr{B}$ occurs*
- *For $\alpha = 1$, phase $\mathscr{C}$ occurs.*
- *For $\alpha \in (1, 2]$ phase $\mathscr{D}$ occurs.*

Phase $\mathscr{A}$ does not occur amongst β-coalescents. However, it is easy to construct Λ-coalescents which are in phase $\mathscr{A}$, such as $\Lambda(dx) = x^2 dx$.

2.4 The Λ-coalescent on a graph

We have mentioned that the Segregated Λ-coalescent is a spatial extension of the Λ-coalescent. One spatial version of the Λ-coalescent has already appeared, in Limic and Sturm (2006), building on the spatial version of Kingman's coalescent from Greven et al. (2005).

Remark 2.4.1 *The model from Limic and Sturm (2006) has been referred to in the literature as 'the Spatial Λ-coalescent'. The dual of the Spatial Λ-Fleming-Viot process, which we will encounter in Section 3.4, has also been refereed to in the literature (see e.g. Barton et al. 2012) as 'the Spatial Λ-coalescent'. These two objects are very different processes.*

Let us give a brief description of the model considered in Limic and Sturm (2006). Let G be a finite graph. Initially, at each vertex $g \in G$ we have a copy of $1^{\mathbb{N}}$ (the partition of $\mathbb{N}$ into singletons). As time passes two mechanisms operate simultaneously. Firstly, each block (from any vertex) has an independent spatial motion, causing it to move between the vertices of G. Additionally, at each vertex $g \in G$ we run mutually independent Λ-coalescents, operating on all the blocks which are currently at the given vertex. Thus, blocks wander freely around G, but are intermittently coagulated with other blocks which happen to be at the same vertex at the same time. Limic and Sturm (2006) obtained conditions analogous to those of Schweinsberg (2000a) for the spatial Λ-coalescent to come down from infinity.

However, Angel et al. (2010) show that, in the same model but with G countably infinite and of bounded degree, the resulting process does not come down from infinity (counting blocks from all sites of G together). Angel et al. (2010) also obtain asymptotic results on the behaviour of their coalescent, in the Kingman and β-coalescent cases, for the behaviour of a large (but finite) number of blocks.

2.5 The Λ-Fleming-Viot process

We mentioned above that there is a duality, due to Bertoin and Le Gall (2003), between the Λ-coalescent and a process known as the Λ-Fleming-Viot process. The Λ-Fleming-Viot process will be particularly important to us in Chapter 3. In fact, the central object considered in this thesis, the Segregated Λ-coalescent, developed out of the Spatial Λ-Fleming-Viot process which is itself a natural extension of the Λ-Fleming-Viot process.

The Spatial Λ-Fleming-Viot process and the Segregated Λ-coalescent are (in essence, at least) spatial equivalents of the Λ-Fleming-Viot process and the Λ-coalescent respectively. However, as we will see in future sections, neither of the spatial processes contains an equivalent of the Kingman component of Λ-coalescents. For this reason we describe the duality between Λ-coalescents and Λ-Fleming-Viot processes without its Kingman component.

Since we have no Kingman component, $\Lambda(\{0\}) = 0$ and in this section we will work with the measure $\nu(dx) = x^{-2}\Lambda(dx)$. The measure ν was introduced by Bertoin and Le Gall (2003) as the natural parametrization for their formulation of the Λ-coalescent.

The Λ-Fleming-Viot process is a measure valued process, which we define by means of a

martingale problem. Let $\mathscr{P}[0,1]$ denote the space of probability measures on $[0,1]$ and let $C(\mathscr{P}[0,1])$ denote the space of continuous real valued functions on $\mathscr{P}[0,1]$. We define the operator G on $C(\mathscr{P}[0,1])$ by

$$G\phi(\rho) = \int_0^1 \rho(dk) \int_0^1 \nu(du) \left[\phi\big((1-u)\rho + u\delta_k\big) - \phi(\rho)\right] \tag{2.5.1}$$

where $\phi \in C(\mathscr{P}[0,1])$. It is easily seen that G is a linear operator and it is shown in Section 5.2 of Bertoin and Le Gall (2003) that the Λ-Fleming-Viot process with generator G can be defined via the following martingale problem.

Definition 2.5.1 (The Λ-Fleming-Viot process) *The Λ-Fleming-Viot process $(\rho_t)_{t\geq 0}$ is the unique càdlàg $\mathscr{P}[0,1]$ valued process such that for all $\phi \in C(\mathscr{P}[0,1])$, the process $t \mapsto \phi(\rho) - \int_0^t G\phi(\rho_s)ds$ is a martingale.*

Examining (2.5.1), the usual heuristic interpretation of the generators of Markov jump processes gives the following description for the evolution of (ρ_t). Let M be a Poisson point process M of rate $dt \otimes \nu(dx)$ with points $(t,x) \in (0,\infty) \times (0,1]$.

- Initially, we have ρ_0 as the uniform measure on $[0,1]$.
- Whenever $(t,x) \in M$, a *reproduction event* occurs. This means that, conditional on ρ_{t-}, we sample an independent random variable k with law ρ_{t-}, and define

$$\rho_t = (1-x)\rho_{t-} + x\delta_k. \tag{2.5.2}$$

- In between reproduction events the process ρ_t is constant.

Of course, this description only makes rigorous sense if $\nu((0,1])$ is finite, but it provides good intuition for the behaviour of the general case. At a reproduction event, we call k the parent site and refer to the atom at δ_k as the children[1] of k.

The formula (2.5.2) will be of great importance in Chapter 3, along with the fact that the Λ-coalescent and the Λ-Fleming-Viot process are indeed dual processes. In words, this duality means that the (forwards in time) evolution of the Λ-Fleming-Viot process completely characterizes the distribution of the (backwards in time) evolution of the Λ-coalescent.

[1]We think of the atom at k as the mass of an infinite set of children, so we use the plural.

We can get a better idea of how the duality works if we recall our comments which followed Definition 2.1.1. Namely, since $\Lambda(\{0\}) = 0$ we can describe the Λ-coalescent in terms of the Poisson point process M as follows.

- Initially, $\Pi_0 = \mathbb{1}^{\mathbb{N}}$, the partition of $\mathbb{N}$ into singletons.
- If (t, x) is a point of M, then a coagulation event occurs at time t. Each block of Π_{t-} chooses, independently and with probability x, whether or not to participate in the coagulation event. All participating blocks are coagulated into a single block, and all non-participating blocks are left untouched, giving Π_t.
- In between coagulation events, Π_t is constant.

Of course, this description also only makes rigorous sense if $\nu((0, 1]) < \infty$.

In very loose terms, the duality is that the reproduction events of (ρ_t) match up to the coagulation events of (Π_t). The branching structure in ρ_t, where the parent point k corresponds to the coagulated block of participating blocks and the children of k correspond to the (uncoagulated) participating blocks.

For completeness we give a precise description of the duality in terms of generators. The classical definition of duality between stochastic processes is a statement concerning only the finite dimensional distributions of the processes involved (see e.g. Chapter 4 of Ethier and Kurtz 1986). Therefore, the intuition conveyed by the above paragraph will appear only in distributional terms.

The Markov generator G^* of the Λ-coalescent (Π_t) can be characterised as follows. For $\psi : \mathcal{P}_p \to \mathbb{R}$ and $\pi \in \mathcal{P}_p$ define

$$G^*\psi(\pi) = \sum_{J \subseteq \{1,\dots,n\}, |J| \geq 2} \lambda_{n,|J|} \left[\psi(m_J \pi) - \psi(\pi)\right] \tag{2.5.3}$$

where $\pi = \{b_i \,;\, i = 1, \dots, n\}$, $m_J\pi$ denotes π with the blocks corresponding to $i \in J$ coagulated.

The summation in (2.5.3) specifies which blocks coagulate, $\lambda_{n,|J|}$ specifies the rates at which the reproduction events corresponding to coagulating the blocks with indices in J occur, and $\psi(m_J\pi) - \psi(\pi)$ is the resulting change to Π_t viewed through the test function ψ. Note that this agrees with Definition 2.1.1.

For each $p \in \mathbb{N}$ and $f \in C([0,1]^p)$ define $\Phi_f : \mathscr{P}([0,1]) \times \mathcal{P}_p \to \mathbb{R}$ by

$$\Phi_f(\rho, \pi) = \int_{[0,1]^n} \rho(dx_1) \dots \rho(dx_n) f\left(Y_\pi(x_1, \dots, x_n)\right)$$

where $n = |\pi|$ is the number of blocks of π and Y_π is defined by $Y_\pi(x_1, \dots, x_n) = (y_1, \dots, y_p)$ where $y_j = x_i$ if (and only if) $i \in b_j$ where $\pi = \{b_j \,;\, j = 1, \dots, p\}$ are the blocks of π ordered by least element.

The duality relationship between Π_t and ρ_t is formally stated as follows.

Lemma 2.5.2 *It holds that*

$$G\Phi_f(\rho, \pi) = G^*\Phi_f(\rho, \pi) \tag{2.5.4}$$

for all $p \in \mathbb{N}$, $f \in C([0,1]^p)$, $\rho \in \mathscr{P}([0,1])$ and $\pi \in P_p$. Hence,

$$\mathbb{E}\left[G\Phi_f(\rho_t, \Pi_0)\right] = \mathbb{E}\left[G^*\Phi_f(\rho_0, \Pi_t)\right] \tag{2.5.5}$$

for all $t \geq 0$.

Remark 2.5.3 *On the left hand side of* (2.5.4) *G acts on Φ_f as a function of its first coordinate whereas on the right hand side G^* acts on Φ_f as a function of its second coordinate.*

PROOF: Equation (2.5.4) can be verified with a direct calculation. Equation (2.5.5) follows from (2.5.4) and the classical theory of martingale problems (which can be found in, for example, Section 4.4 of Ethier and Kurtz 1986). Alternatively, see Lemma 4 of Bertoin and Le Gall (2003). ■

Chapter 3

The Spatial Λ-Fleming-Viot process

The Λ-coalescent generalized Kingman's coalescent by allowing mergers involving more than two (blocks of) particles. In another vein, efforts were made to further generalize population models and incorporate the effects of the geographical space. For our purposes the transition towards spatial models begins with the Λ-Fleming-Viot process of Section 2.5.

The Spatial Λ-Fleming-Viot process (SΛFV) first appeared in Etheridge (2008), as a spatial extension of the Λ-Fleming-Viot process. The SΛFV is an infinite system (one at each site of $\mathbb{R}^d$) of interacting Λ-Fleming-Viot processes. For the duration of this section we fix $d \in \mathbb{N}$.

The terminology 'the SΛFV process' is usually reserved for the process considered in Barton et al. (2010a), but both Etheridge (2008) and Barton et al. (2010a) stress that the SΛFV is only one illustration of a much more general theme, namely population models in a spatial continuum with reproduction controlled by a Poisson point process. We refer to such a process as 'a SΛFV process'. The variety of possibilities is demonstrated by Berestycki et al. (2009), Barton et al. (2010b) and Etheridge and Véber (2012), each of which features a different modification of the original SΛFV process.

In this chapter we work with a modification of the original SΛFV process. Our version will be more than sufficient to illustrate the features of the process that we wish to describe. The difference is described in Remark 3.2.2 and is very minor, so in this thesis we choose to refer to the process described by Definition 3.2.1 as 'the SΛFV process'.

The population described by the SΛFV process lives in $\mathbb{R}^d$ and the genetic types of individuals within the population are drawn from the compact metric space $\mathcal{K}$. Multiple individuals (in a sense described below) occupy each point of $\mathbb{R}^d$.

3.1 The state space

The SΛFV process has a non-trivial state space Ξ which we will now describe. Let $\mathscr{P}(\mathcal{K})$ be the space of probability measures on $\mathcal{K}$, equipped with the weak topology. Let $\widetilde{\Xi}$ be the space of all measurable maps from $\mathbb{R}^d$ to $\mathscr{P}(\mathcal{K})$. We set

$$\rho_1 \sim \rho_2 \text{ iff } \{x \in \mathbb{R}^d ;\, \rho_1(x) \neq \rho_2(x)\} \text{ is Lebesgue null,} \tag{3.1.1}$$

and define Ξ to be the quotient of $\widetilde{\Xi}$ under $\sim$.

The space Ξ is the state space of the SΛFV process, with the topology defined in Appendix C.1. Without going into details, in Appendix C.1 we see that Ξ is a subset of the unit ball of a suitable Banach space, which is itself equipped with a weak-star topology; the space Ξ is given the subspace topology from this embedding. This topology originates from Evans (1998) and is the same as was used by Barton et al. (2010a).

It is shown in Appendix C.1 that Ξ is a compact metrizable space which comes equipped with a natural set of test functions (i.e. appropriate for use with the generator of the SΛFV process). The formal statements of existence and duality of the SΛFV process, which appear in a heuristic form in Sections 3.2 and Sections 3.4, are reliant on a precise description of the topology on Ξ because they use the Markov generator and semigroup of the SΛFV process. We give the formal statements in Appendices C.2 and C.3 respectively

3.2 Definition

The SΛFV process is a Ξ valued stochastic process. The parameters of the SΛFV process are a measure $\mu(dr)$ on $(0, \infty)$ and a family of probability measures $\nu_r(du)$ on $[0, 1]$ such that $\mu(dr)\nu_r(du)$ is a measure on $(0, \infty) \times [0, 1]$. We require that

$$\int_0^\infty \int_0^1 u r^d \, \nu_r(du)\mu(dr) < \infty. \tag{3.2.1}$$

We will discuss what this condition means for the SΛFV process (and why it is needed) after the definition.

The SΛFV process has the following description in terms of a Poisson point process. Let M be a Poisson point process with points (t, y, r, u) in $(0, \infty) \times \mathbb{R}^d \times (0, \infty) \times [0, 1]$ and rate

$$dt \otimes dy \otimes \mu(dr)\nu_r(du).$$

For $r > 0$ and $y \in \mathbb{R}^d$, recall that $B_r(y) = \{z \in \mathbb{R}^d ;\, |y - z| < r\}$.

Definition 3.2.1 (The SΛFV process) *From some initial state $\rho_0 \in \Xi$ the evolution is specified as follows.*

- *Whenever $(t, y, r, u) \in M$, sample k according to $\rho_{t-}(y)$. Set*

$$\rho_t(x) = (1-u)\rho_{t-}(x) + u\delta_k \tag{3.2.2}$$

  for all $x \in B_r(y)$.

- *For each $x \in \mathbb{R}^d$, in between the jump times caused by reproduction events (as above) the process $\rho_t(x)$ is constant.*

We refer to (t, y, r, u) as a reproduction event occurring at time t about y with radius r and killing proportion u. We refer to y as the parent site and k or δ_k as the parent type. We say the region $B_r(y)$ was affected (or hit) by the reproduction event.

Definition 3.2.1 only makes formal sense when

$$\int_0^\infty (1 \vee r^d)\mu(dr) < \infty,$$

in which case reproduction events affect each bounded region of $\mathbb{R}^d$ at finite rate. A mathematically proper definition of the SΛFV process requires a significant amount of notation and is given (using its generator) in Appendix C.2.

The measure $\rho_t(x)$ should be thought of as representing the local distribution of genetic types. The SΛFV process is not concerned with the quantity of individuals present in a spatial location, only the relative frequencies of their genetic types.

From equation (3.2.2) we can see the SΛFV process is an infinite system of interacting Λ-Fleming-Viot processes. In a single Λ-Fleming-Viot process the parent was always sampled from within that process, whereas in Definition 3.2.1 the parent can be sampled from nearby sites.

Remark 3.2.2 *In Definition 3.2.1 we sample the parent type k from $\rho(y)$, where y is the center of the ball in which the reproduction event takes place. In Barton et al. (2010a), the parent is sampled from $\rho(Y)$ where $Y \in B_r(y)$ is sampled (independently of all else) upon each reproduction event from the uniform distribution on $B_r(y)$. Barton et al. (2010a) defined the SΛFV process on a two dimensional torus, but there is no material difference to the definition of the process when this torus is replaced by $\mathbb{R}^d$.*

Definition 3.2.1 is the same version of the SΛFV process as was considered in Saadi (2011).

3.3 The rate of reproduction

In Definition 3.2.1 it is not immediately clear what it means to be an 'individual living within the population'. In this section we seek to describe what it means to be such an individual in such a way as the ancestral lineage of each individual is a finite rate jump process.

Note that the rate at which a single point $x \in \mathbb{R}^d$ is hit by reproduction events in the SΛFV process is

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}\{x \in B_r(y)\}\, \nu_r(du)\mu(dr)dy = C_d \int_0^\infty r^d \mu(dr), \tag{3.3.1}$$

where C_d denotes the volume of a d dimensional unit ball. The term $\int_0^\infty r^d \mu(dr)$ is potentially infinite, despite (3.2.1).

Each point $x \in \mathbb{R}^d$ carries one of the (interacting) Λ-Fleming-Viot processes, so from mathematical point of view it is natural to think of each point as containing an infinity of individuals. To be precise, a point $x \in \mathbb{R}^d$ is associated to a probability measure $\rho_t(x)$ on $\mathcal{K}$ which represents the local distribution of genetic types.

We can embed a genealogy into the SΛFV process as follows. At each site x we have a local neighbourhood containing uncountably many individuals indexed by $[0, 1]$. The measure $\rho_t(x)$ specifies the distribution of types of these individuals, where each individual has a single genetic type drawn from $\mathcal{K}$. At each reproduction event (t, y, r, u) the type k of the parent individual is sampled from $\rho_{t-}(y)$. The variable u specifies what proportion of the individuals at x (for each $x \in B_r(y)$) are replaced by the parent type k. We refer to the replacements as the children (of the parent with type k) who are born at this reproduction event.

Remark 3.3.1 *By* (3.2.2)*, in the reproduction event caused by* $(t, y, r, u) \in M$*, each genetic type originally present in* $\rho_t(x)$ *(for* $x \in B_r(y)$*) has a fraction* u *of itself replaced.*

Consider tracing back the ancestral lineage of a single individual in the dynamics described above. If the lineage is at the point x at time t and was affected by a reproduction

event with parameters (t, y, r, u), which means that $x \in B_r(y)$, then with probability $1-u$ the lineage stayed put at the point x and with probability u it came from the point y at time $t-$.

A single lineage has its position affected by reproduction events with rate

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}\{0 \in B_r(x)\} u \nu_r(du)\mu(dr)dx = C_d \int_0^\infty \int_0^1 r^d u \, \nu_r(du)\mu(dr), \tag{3.3.2}$$

where C_d denotes the volume of the unit ball in $\mathbb{R}^d$. Note that, by the translation invariance of Definition 3.2.1, the displacement by which the lineage changes its position (when it does actually move) is independent of the location of the lineage before the jump. Hence, a single lineage follows a compound Poisson process. The ancestral lineages of two individuals are not independent of one another, even before coalescence, since a large enough reproduction event has a positive probability of affecting the location of both lineages. Thus the full system of ancestral lineages is a dependent system of coalescing compound Poisson processes.

We refer to the above description of the SΛFV process as the *notional formulation* of the process. Under these heuristics, each individual has a single genetic type and lives for an exponential $C_d \int_0^\infty \int_0^1 r^d u \, \nu_r(du)\mu(dr)$ time. In the notional description there are uncountably many individuals associated to each site (and there are uncountably many sites). It is sensible to expect that the existence of the SΛFV process is a delicate question.

3.4 The dual of the SΛFV process

Since the SΛFV process is a spatial extension of the Λ-Fleming-Viot process and the Λ-Fleming-Viot process is itself dual to the Λ-coalescent, it is natural to expect the dual of the SΛFV process to behave like a spatial extension of the Λ-coalescent. We will shortly see that this does indeed turn out to be the case, but in Section 3.7 we will see that the framework of the SΛFV process imposes several restrictions on the behaviour of the coalescent, such as not coming down from infinity. In fact, in the dual of the SΛFV process we only see some of the equivalent behaviour to the Λ-coalescent.

Barton et al. (2010a) established the existence of the SΛFV process[1] via the following dual process, using an adaptation of a method of Evans. Evans (1998) is a general tool

[1]Modulo the modifications described in Remark 3.2.2, which have no significant effect on the analysis.

for constructing processes which are dual to a specified system of ancestral lineages. From our own point of view, we have already seen a definition of the SΛFV process but we have also seen that the notional formulation of Section 3.3 suggests a way to define its ancestral lineages.

We now give a full description of dual process of the SΛFV process. Note that, since we dealing with a dual process, we describe the motion of the lineages with the direction of time reversed from that of the SΛFV process.

Definition 3.4.1 (The dual of the SΛFV process) *Let M be a Poisson point process with points $(t, y, r, u) \in [0, \infty) \times \mathbb{R}^d \times (0, \infty) \times [0, 1]$ of rate*

$$dt \otimes dx \otimes \mu(dr)\nu_r(du).$$

Fix $N \in \mathbb{N}$. We will define a system of N particles (which we refer to as ancestral lineages) moving around $\mathbb{R}^d$. Let B_t^n denote the position of the n^{th} lineage at time $t \geq 0$. We will also need a process $\sim_t$, taking values in the equivalence relations on $\{1, \ldots, N\}$. Fix some initial state $(B_0^n)_{n=1}^N \in \mathbb{R}^d$ and define the evolution as follows.

- *If $(t, y, r, u) \in M$, let A denote the set of equivalence classes of $\sim_{t-}$. For each $a \in A$, write $B_t^a = B_t^{a_0}$ where $a_0 \in a$, and note this does not depend on the choice of $a_0 \in a$.*

  For each $a \in A$, if $B_{t-}^a \in B_r(y)$ then, independently of all else, sample a Bernoulli random variable with success probability u and on a success set $B_t^n = y$ for all $n \in a$. Coalesce the blocks $a \in A$ which saw a success to obtain $\sim_t$ from $\sim_{t-}$. On a failure do nothing.

- *In between reproduction events causing jumps as above, the lineages do not move and $t \mapsto \sim_t$ is constant.*

We say that a lineage was affected by a reproduction event if the reproduction caused the lineage to change its location. (Note that a reproduction event which affects the area $B_r(y)$ will only affect some of the lineages which were located in $B_r(y)$ when the event happened.)

The equivalence relations keep track of blocks of coalesced lineages. If at any point in time two lineages B_t^m and B_t^n are involved in the same reproduction event (i.e. the Bernoulli random variable described above is sampled for the relevant blocks and is a success in each

case), they coalesce and remain together for all further time. Having realized that, in order to understand Definition 3.4.1 we need only to understand the motion of a single lineage; we already did so in Section 3.3 and the description there agrees with Definition 3.4.1.

Since $\mu(dr)\nu_r(du)$ is potentially an infinite measure we must check that (3.2.1) implies the system in Definition 3.4.1 is well defined. First note that, under Definition 3.4.1, if two lineages are at separation $z \in \mathbb{R}^d \setminus \{0\}$ then they coalesce at rate

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}\{0, z \in B_r(y)\} y^2 \, \nu_r(du)\mu(dr)dy. \tag{3.4.1}$$

Performing the integration with respect to y and using the bound $L_r(0,z) \leq C_d r^d$, where $L_r(x,y)$ denotes the volume of $B_r(x) \cap B_r(y)$, we get

$$(3.4.1) \leq C_d \int_0^\infty \int_0^1 u^2 r^d \, \nu_r(du)\mu(dr).$$

Since $u^2 \leq u$ the condition (3.2.1) is more than sufficient to guarantee (3.4.1) is finite. A similar calculation can be done for multiple coalescence events.

Secondly, under Definition 3.4.1, a single lineage jumps at rate

$$\int_0^\infty \int_0^1 \mathbb{1}\{0 \in B_r(x)\} u\nu_r(du)\mu(dr) \, dx. \tag{3.4.2}$$

Thus a lineage would correspond to a well defined Lévy process $t \mapsto \sum_{(s,\Delta x)\in \widetilde{M}} \mathbb{1}\{s < t\}\Delta x$ (for some suitable Poisson point process $\widetilde{M}$) if

$$\int_{\mathbb{R}^d} (1 \wedge |x|^2) \left(\int_0^\infty \int_0^1 \mathbb{1}\{0 \in B_r(x)\} u\nu_r(du)\mu(dr) \right) dx \tag{3.4.3}$$

was finite. As we saw in (3.3.2), equation (3.2.1) gives us something much stronger; it says that (3.4.2) is finite, so that the lineages jump at a finite rate. These two checks, whilst not a formal proof, establish beyond reasonable doubt that the dual system is well defined.

In view of (3.4.3) one might suspect that (3.2.1) is in fact too strong and that, with more care, the SΛFV process could exist with only a weakened version of (3.2.1). This leads us to our next section, but first let us mention that a precise statement of the duality between the SΛFV process and the system of ancestral lineages defined by Definition 3.4.1 is given in Appendix C.3.

3.5 The limitation on the rate of reproduction

Let us briefly take stock. We defined the SΛFV process in Definition 3.2.1 and noted that the condition (3.2.1) imposed a limitation on the rate at which reproduction events occurred

within the process. In Section 3.4 we saw a notional formulation of the process, which led to a corresponding dual process of coalescing compound Poisson processes, defined in Definition 3.4.1. We saw that (3.2.1) made the ancestral lineages jump at finite rate, which was more than sufficient to show that the dual process was well defined.

As one might imagine from Section 3.4, abandoning the condition imposed by (3.2.1) would correspond to replacing the dual with a system of coalescing Lévy processes. We run into difficulty trying to represent all the resulting lineages with dynamics driven by a Poisson point process. This is most easily seen if we attempt to write down the generator G of the SΛFV process.

Let us do so on the test function $I_1 : \Xi \to \mathbb{R}$ by

$$I_1(\rho) = \int_{\mathbb{R}^d} \psi(x) \int_{\mathcal{K}} \chi(k)\, \rho(x)(dk)\, dx$$

where $\psi : \mathbb{R}^d \to \mathbb{R}$ is integrable and $\chi : \mathcal{K} \to \mathbb{R}$ is continuous (which implies that χ is bounded by compactness of $\mathcal{K}$). It can be seen from results in Appendix C.1 that I_1 is continuous on Ξ. Definition 3.2.1 suggests that the result will be

$$GI_1(\rho) = \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_0^1 \nu_r(du) \int_{\mathcal{K}} \rho(y)(dk) \int_{\mathbb{R}^d} dx \\ \Big[\psi(x)\mathbb{1}\big\{x \in B_r(y)\big\}\Big(\langle \chi, (1-u)\rho(x) + u\delta_k\rangle - \langle \chi, \rho(x)\rangle \Big)\Big].$$

This can be checked against our definition of the generator in Appendix C.2. A little rearrangement leads us to

$$GI_1(\rho) = \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_0^1 \nu_r(du) \int_{\mathbb{R}^d} dx \Big[\psi(x)\mathbb{1}\{x \in B_r(y)\} u\Big(\langle \chi, \rho(y)\rangle - \langle \chi, \rho(x)\rangle \Big)\Big] \tag{3.5.1}$$

The best bound on the terms in large round brackets is in general $2||\chi||_\infty$. Since $||\phi||_1 < \infty$ and $\int_{\mathbb{R}^d} \mathbb{1}\{x \in B_r(y)\}dy = C_d r^d$ we have

$$|GI_1(\rho)| \le 2||\chi||_\infty ||\psi||_1 \int_0^\infty \mu(dr) \int_0^1 \nu_r(du)\, u r^d$$

From this it is apparent that (3.2.1) is precisely the right bound for GI_1 to be well defined.

Note that the test function I_1 is a very natural test function to use on $\rho \in \Xi$. It integrates over both space and genetic type, with the function ψ weighting the spatial locations and the function χ weighting the genetic types.

Remark 3.5.1 *In Appendices C.1 and C.2 we establish a natural set $\mathscr{I}$ of continuous functions on Ξ on which to define the generator of the SΛFV process. A similar calculation to the above can be carried out for any $I \in \mathscr{I}$.*

In Section 2.5 we exhibited suitable test functions to use for the generator of the Λ-Fleming-Viot process. These were essentially I_1 without ψ and without the integral over $\mathbb{R}^d$, in other words they were I_1 without its spatial element. In the generator of the Λ-Fleming-Viot process we do not need a condition like (3.2.1) *because without space the parent is selected from 'the same site' as it reproduces into; there is no equivalent of the $\mathbb{1}\{x \in B_r(y)\}$ which appeared in* (3.5.1).

We end this section with the comment that it is known that some versions of the SΛFV exist with a corresponding notional formulation in which the ancestral lineages are infinite rate pure jump Levy processes (Etheridge and Véber 2010). Since the Poisson point process representation breaks down it is not known how to characterise these processes forwards in time.

Expressing the SΛFV process as driven by reproduction events taking place in finite regions, where each location is affected equally (at least, in distribution), puts a limitation on the type of interaction between the ancestral lineages. As a consequence the Poisson point process driven SΛFV cannot support a sufficiently complex dependency between the ancestral lineages as would be needed for a system of coalescing infinite rate Lévy processes.

3.6 Review of the SΛFV literature

In this section we take a break from our development of the SΛFV process and give a brief overview of its current literature.

1. **Etheridge (2008)** gives a survey of models used in modern population genetics. The article contains a short section introducing the SΛFV and a list of suggestions for generalizations of the process.

2. **Berestycki, Etheridge, and Hutzenthaler (2009)** consider a process which is intended to be a particle system approximation to the SΛFV process, with only finitely many individuals in any bounded region. In contrast to most finite population models the individuals in their model exhibit correlated reproduction. The article is concerned

mostly with discussing the long term survival and ergodicity of such models. Proof that these models really do converge to the SΛFV process is to appear in Etheridge and Kurtz (2012).

3. **Barton, Etheridge, and Véber (2010a)** gave the first construction of the SΛFV process. Their construction is, as we have already mentioned, an adaptation of a method of Evans (1998) but the primary aim of the article is to investigate the asymptotic properties of the dual. Under different scaling limits they obtain a Kingman coalescent, a general Λ-coalescent and a system of coalescing Brownian motions with a non-local coalescence mechanism.

4. **Barton, Kelleher, and Etheridge (2010b)** investigate one of the biological effects that the SΛFV process was designed to model. They consider a version of the process in $\mathbb{R}^2$ where each reproduction event affects all of $\mathbb{R}^2$ but the proportion u of individuals replaced at a site y is the value of a Gaussian function centred about the parent location. A simulation package (which is used in the paper) can be found at `http://homepages.ed.ac.uk/jkellehe/qps.php`. Using simulations, they investigate the extent to which large reproduction events cause 'bottlenecks' in the genealogy of the population, leading to long range (in spatial terms) correlation between the genetic types of individuals.

5. **Berestycki, Etheridge, and Véber (2012c)** looks at scaling limits of two versions of the SΛFV process. Firstly, a case where the ancestral lineages are required to have finite variance and are rescaled to Brownian motion. Secondly, a version with a heavy tailed distribution on the radii of reproduction events where the ancestral lineages are rescaling to α-stable Lévy processes. In the first case, in all but dimension one, two independent Brownian motions don't meet and this leads to a spatially homogeneous and deterministic limit.

   One might initially suspect that the same would hold in the second case, at least in all spatial dimensions where two independent α-stable processes do not meet. However, in the second case it is shown that in *all* dimensions the large extinction/recolonisation events lead to non-trivial correlations between the limiting genetic types at different sites. The cause is the correlation between the ancestral lineages in the dual of the

SΛFV process which, in the limit, mean that any finite sample of individuals have their most recent common ancestor only finitely long ago.

6. **Véber and Wakolbinger (2012)** is work in progress and is intended to provide a lookdown construction of the SΛFV process, in the style of **?**. They also express the SΛFV process as a measure valued process, in particular as an $\mathcal{M}(\mathbb{R}^d \times \mathcal{K})$ valued process; this is achieved via establishing a natural homeomorphism between Ξ and $\mathcal{M}(\mathbb{R}^d \times \mathcal{K})$.

7. **Etheridge and Véber (2012)** extend the model of Barton et al. (2010a) to diploid populations and incorporate a natural recombination mechanism. They study the correlations between the genetic types of individuals at different points in space, giving a potential means of inferring from data the existence of large extinction/recolonization events in a populations history.

8. **Véber and Yu (2012)** is work in progress and concerns connections between the SΛFV process and the FKPP equation.

9. **Barton, Etheridge, and Véber (2012)** is a survey article which discusses many of the articles listed above in greater depth.

3.7 The dual of the SΛFV process as a coalescent

In the section we examine the dual of the SΛFV process in a similar fashion to that in which we examined the Λ-coalescent in Section 2.3. In the interests of brevity, we consider a slightly simplified version of the SΛFV process; in particular we assume that ν_r does not vary with r, so as we can write

$$\mu(dr)\nu_r(du) = \mu(dr) \otimes \nu(du) \tag{3.7.1}$$

as a product measure. Note that, with (3.7.1), the condition (3.2.1) for existence of the SΛFV process reduces to

$$\int_0^\infty r^d \mu(dr) < \infty. \tag{3.7.2}$$

Following our discussion in Section 3.3, we know that (3.7.2) means that the rate of reproduction events affecting a single point is finite. However, it is easy to see that the rate of reproduction event affecting a bounded (non-null) region of $\mathbb{R}^d$ may be infinite.

In Definition 3.4.1 we defined a dual process to the SΛFV which was composed of $N \in \mathbb{N}$ particles. With a projective limit we can define an analogous system for countably many particles, starting with one particle at each point of a dense subset of $\mathbb{R}^d$. Let $\mathbb{Q}^d$ denote the (countable) set of points of $\mathbb{R}^d$ with rational coordinates and let $(B^q)_{q\in\mathbb{Q}^d}$ denote the corresponding system defined by a projective limit of Definition 3.4.1.

We say B^q and B^p are in the same block at time $t > 0$ if B^q and B^p were both affected by a common reproduction event during $(0, t]$. For the rest of this section, we write $p \sim_t q$ when B^q and B^p are in the same block at time t and we say that B^q and B^p are coalesced lineages at time t. Define

$$\Delta_t = \mathbb{Q}^d / \sim_t \tag{3.7.3}$$

as the quotient of $\mathbb{Q}^d$ under the equivalence relation $\sim_t$, so that elements of Δ_t are blocks of coalesced lineages, with lineages labelled according to their initial locations. In the spirit of Definition 2.2.1, we say that (Δ_t) comes down from infinity if $\mathbb{P}[\forall t > 0, |\Delta_t| < \infty] = 1$.

Since each reproduction event affects a subset of $\mathbb{R}^d$ which has positive Lebesgue measure, blocks of coalesced individuals are either singletons or contain an infinity of points.

Our first observation is the following.

Lemma 3.7.1 *Let $t > 0$ and let A be a non-empty open subset of $\mathbb{R}^d$. Then, with positive probability, there is a set $A' \subseteq A$ such that A' has positive Lebesgue measure and none of the lineages B^q with $q \in A'$ were affected by a reproduction event during time $[0, t]$.*

Proof: Define $g(x) = \inf\{s > 0\,;\, (s, y, r, u) \in M, x \in B_r(y)\}$. Note that $B^q(t) = B^q(0)$ for all $t \leq g(x)$, since the first reproduction event to affect B^q happens after time $g(x)$. Note that

$$\mathbb{P}[g(x) > t] = \mathbb{P}[g(0) > t] = \exp\left[-tC_d\left(\int_0^1 u\,\nu(du)\right)\left(\int_0^\infty r^d\,\mu(dr)\right)\right] > 0$$

Let $A \subseteq \mathbb{R}^d$ be non-empty and open, and set $A' = \{x \in A\,;\, g(x) \geq t\}$. By Fubini's theorem[2],

$$\mathbb{E}\left[A'\right] = \mathbb{E}\left[\int_A \mathbb{1}\{x \in A'\}dx\right] = \int_A \mathbb{E}\left[\mathbb{1}\{x \in A'\}\right] dx = \left(\int_A dx\right)\mathbb{P}[g(0) > t] > 0.$$

Hence, there is positive probability that A' has positive Lebesgue measure. ■

When A' has positive Lebesgue measure it is necessarily uncountable. By Lemma 3.7.1, even if we restrict ourselves to a bounded region A of $\mathbb{R}^d$, (Δ_t) does not come down from

[2]For brevity, we omit a check of the measurability of $(\omega, x) \to g(x)(\omega)$.

infinity; at any time $t > 0$, the initial condition (i.e. the partition into singletons) can be found with positive probability on infinitely many $q \in A \cap \mathbb{Q}^d$. It is the fact that lineages only jump at finite rate which causes (Δ_t) to not come down from infinity.

Note that, unlike the Λ-coalescent, it is possible to have $\mathbb{P}\left[|A \cap \Delta_t| < \infty\right] \in (0, 1)$ even if $A \subseteq \mathbb{R}^d$ is infinite (e.g. if A was contained in a reproduction event for which $u = 1$). We will see that the equivalent of this also occurs in the Segregated Λ-coalescent.

There is another unsatisfactory aspect to our treatment of the SΛFV dual as a coalescent. Because we had to sample a uniform random variable for each lineage on each of its (potential) jumps, we are only able to define the dual process for a countable number of lineages. In our notional formulation we had infinitely many individuals at every point of $\mathbb{R}^d$, meaning that the dual processes is not able to capture the full picture of the behaviour of the ancestral lineages.

It is natural to ask whether it is possible to give a more precise formulation of the SΛFV process in which it is possible to construct the dual of the SΛFV process in such a way as, in some sense, we see all the ancestral lineages of the process at once. The question seems difficult and we constrain ourselves here to noting that such a construction would require a different state space for the SΛFV process. This is because for each $\rho \in \Xi$, $\rho(x)$ is only defined for x outside of a null set $\mathcal{N}$ of $x \in \mathbb{R}^d$ (due to (3.1.1)) and the null set $\mathcal{N}$ potentially contains an infinity of ancestral lineages.

We will see in Chapter 5 that, at least for the Segregated Λ-coalescent, it is possible for the set of singletons to occupy a null but non-empty proportion of the geographical space. Thus the issue discussed in the above paragraph is a serious one with regards to analysing coalescent processes in spatial continua.

Our approach to the issues discussed above is to ask: Is it possible to define a (not the) SΛFV process such that:

1. The population lives in a spatial continuum.

2. There is a mathematically precise definition of what it means to be an individual in the population, such that ancestral lineages of all individuals within the process can be simultaneously defined.

3. The model has a dual process in the form of a coalescent made up of ancestral lineages.

4. This coalescent is able to come down from infinity.

In the next chapter we will exhibit a new model called the Segregated Λ-Fleming-Viot process which has all the properties listed above. The dual of the Segregated Λ-Fleming-Viot process is the Segregated Λ-coalescent.

3.8 Uniform killing

Consider the SΛFV process in the case that

$$\nu_r = \delta_{\{1\}}, \tag{3.8.1}$$

the pointmass at one. The effect of this is that, when a site x is affected by a reproduction event caused by $(t, y, r, 1) \in M$, *all* the individuals at x are replaced with individuals of the same type as the parent type. As a consequence, providing we start from $\rho \in \Xi$ such that for all x, $\rho(x) = \delta_k$ for some $k \in \mathcal{K}$, we may think of each point as carrying only a single individual.

In terms of the dual process, it means that if a lineage B^a is inside a region $B_r(y)$ at time t in which a reproduction event occurs, with probability one that lineage will moved to the parent location. Thus, the 'delayed' aspect of the dual system of coalescing compound Poisson processes is removed by (3.8.1). Under (3.8.1), (3.2.1) reduces to $\int_0^\infty r^d \mu(dr) < \infty$ which is still equivalent to saying that the ancestral lineages jump at finite rate. We refer to the SΛFV process with (3.8.1) as the SΛFV process with *uniform killing*.

At first sight (3.8.1) looks to be a serious simplification of the SΛFV process but it turns out that many interesting features of the process can still be seen. For example, there can be multiple different types of individual in any subset of $\mathbb{R}^d$ and the reproduction mechanism still acts in such a way as the genetic types found at a pair of sites $x, y \in \mathbb{R}^d$ are more strongly correlated when x and y are close together.

In fact, with uniform killing the SΛFV process behaves similarly, in a qualitative sense, to a *single* Λ-Fleming-Viot process. If we imagine for a moment that we are running the SΛFV process only on some bounded subset of $\mathbb{R}^d$, or on a sphere so as we don't have to think about what happens at the edges, then what we see is the following. At a reproduction event, the parent individual is selected uniformly in space which means that the genetic type of the parent is sampled uniformly from the whole population, as in the Λ-Fleming-Viot

process. The offspring of this individual then replace a non-trivial proportion of the current population. The difference to the Λ-Fleming-Viot process is that individuals which are replaced are close, in the geographical sense, to the parent; whereas in the Λ-Fleming-Viot process there was no concept of 'distance between individuals' and the individuals replaced were (effectively, at least) sampled uniformly from the whole population.

Chapter 4

The Segregated Λ-coalescent

In this chapter we rigorously define the Segregated Λ-coalescent, which will be the central object of interest for the remainder of the thesis. In doing so we seek to capture mathematically the intuition which was conveyed in Section 1.2 by Definition 1.2.1, which we repeat here for convenience. Recall from Section that M was a Poisson point process with points $(t, w, p) \in (-\infty, \infty) \times W_* \times K$ with rate given by

$$dt \otimes \mathcal{R}(dw)\mathcal{U}_w(dp).$$

Definition 1.2.1 *We define a stochastic flow $(X_{s,t})_{-\infty<s<t<\infty}$ on K as follows.*

- *Whenever $(t, w, p) \in M$, any particles in the flow which are in K_w at time $t-$ jump to p at time t.*
- *The position of each particle is constant in between its jumps.*

A graphical demonstration of the definition was given in Figure 1.1.

Under Definition 1.2.1, a single point of K is affected by reproduction events at rate $\sum_0^\infty r_n$. Following our discussion of the SΛFV process in Section 3.5, we might wonder if the condition $\sum_n r_n < \infty$ was necessary for our models existence. This is not so, in fact in Section 4.2 we give a direct definition of $(X_{s,t})$ without placing any assumptions on $(r_n) \subseteq [0, \infty)$ or $\mathcal{S} \in \mathbb{N} \setminus \{1\}$.

4.1 Segregated spaces

We begin the construction of our model with a description of the general form of its geographical space. Recall that $\mathcal{S} \in \{2, 3, 4, \ldots\}$ and

$$S = \{1, 2, \ldots, \mathcal{S}\}. \tag{4.1.1}$$

Recall that W_n is the set of words $w = w_1 w_2 \dots w_n$ of length n with letters $w_i \in S$ and $W_* = \bigcup_{n=0}^{\infty} W_n$. Recall also that for each $w = w_1 w_2 \dots w_n \in W_*$ we write $|w| = n$. If $w = w_1 \dots w_n$ and $i \in S$ then we set $wi = w_1 \dots w_n i \in W_{n+1}$.

We define the following (non-standard) structure, which is the general form of the geographical space in our model. If (K, D_K) is a metric space then define $\mathrm{diam}(A) = \sup\{D_K(x,y)\,;\, x,y \in A\}$ for all $A \subseteq K$. Recall that $A \uplus B$ denotes the disjoint union of A and B.

Definition 4.1.1 *Let (K, D_K) be a complete metric space, equipped with a family of non-empty measurable subsets $(K_w)_{w \in W_*}$ and a probability measure λ. We say K is a segregated space if it satisfies:*

[$\mathscr{K}1$] $K = K_\emptyset$ and for all $w \in W_*$, $K_w = \biguplus_{i \in S} K_{wi}$.

[$\mathscr{K}2$] There exists a sequence $(L_n) \subseteq (0, \infty)$ such that $L_n \to 0$ and $\mathrm{diam}(K_w) \leq L_{|w|}$.

[$\mathscr{K}3$] If $|w| = |w'|$ then $\lambda(K_w) = \lambda(K_{w'})$.

In the same spirit as (1.2.1), [$\mathscr{K}1$] implies that for all $n \in \mathbb{N}$,

$$K = \biguplus_{w \in W_n} K_w. \tag{4.1.2}$$

We will use [$\mathscr{K}1$] so frequently that it would be impractical to reference it on every application. However, we will not use the other conditions without explicitly saying so. To reconcile with our previous definition, for a general segregated space, K_w is said to be a complex of K with level $|w|$. We extend the definition of subcomplex in the obvious manner.

The point of [$\mathscr{K}2$] is as follows. Suppose $(w(n))$ is a sequence in W_* such that $|w(n)| \to \infty$ and $K_{w(n+1)} \subseteq K_{w(n)}$, and suppose (x_n) is a sequence in K such that $x_n \in K_{w(n)}$. Then [$\mathscr{K}2$] implies that (x_n) is Cauchy, and we can use the completeness of K to infer existence of a limit $x_n \to x \in K$. Note, however, that since $K_{w(n)}$ may not be closed we cannot deduce that $x \in \cap_n K_{w(n)}$; we might have $\cap_n K_{w(n)} = \emptyset$.

Due to [$\mathscr{K}3$], it is natural to think of λ as a uniform measure on K. The measure λ plays no part in the construction of the flow X, but it will provide the distinction between the dust and the non-singleton blocks, as the following result suggests.

Lemma 4.1.2 *For all $w \in W_*$, $\lambda(K_w) > 0$. For all $x \in K$, $\lambda(\{x\}) = 0$.*

PROOF: Let $w \in W_n$. By (4.1.2), for all $n \in \mathbb{N}$, $K = \biguplus_{v \in W_n} K_v$ and by [$\mathscr{K}$3],

$$\lambda(K) = \sum_{v \in W_n} \lambda(K_v) = \mathcal{S}^n \lambda(K_w). \tag{4.1.3}$$

Since $\lambda(K) = 1$, $\lambda(K_w) > 0$.

Suppose that, for some $x \in K$, $\lambda(\{x\}) = \delta > 0$. Then for all $n \in \mathbb{N}$, $x \in K_w$ for some $w \in W_n$. By (4.1.3), $\lambda(K) \geq \mathcal{S}^n \delta$. Since $n \in \mathbb{N}$ was arbitrary this is a contradiction, so $\lambda(\{x\}) = 0$. ■

Example 4.1.3 *The $\mathcal{S}$-part Cantor set, as defined in Section 1.2, with its Bernoulli measure, is a segregated space.*

More generally, recall that an iterated function system (IFS) is a non-empty compact metric space (M, m) equipped with a finite family $(F_i)_{i=1}^n$ of bi-Lipschitz contractions $F_i : M \to M$. A well known theorem of Hutchinson (1981) states that there is a unique non-empty compact subset $\mathcal{A}$ of M, called the attractor, such that $\mathcal{A} = \bigcup_i F_i(\mathcal{A})$. The attractor $\mathcal{A}$ of an IFS which satisfies $\mathcal{A} = \biguplus_i F_i(\mathcal{A})$, equipped with the corresponding Bernoulli measure and with $K_{i_1 \ldots i_n} = F_{i_1} \circ \ldots \circ F_{i_n}(\mathcal{A})$, is a completely segregated space.

Example 4.1.4 *Many examples of segregated spaces can be constructed from iterated function systems where the n-level complexes overlap slightly, by arbitrarily choosing which subcomplexes any overlap should belong to. For example, (take λ to be Lebesgue measure on $[0, 1]$ and) note*

$$[0, 1] = [0, 1/2] \cup [1/2, 1] = [0, 1/4] \cup [1/4, 1/2] \cup [1/2, 3/4] \cup [3/4, 1] = \ldots$$

We could decide that whenever we see an overlap between complexes the point belongs to the complex which is closer to the origin. Thus we would have $[0, 1]$ decomposed into a disjoint subcomplexes as follows:

$$[0, 1] = [0, 1/2] \uplus (1/2, 1] = [0, 1/4] \uplus (1/4, 1/2] \uplus (1/2, 3/4] \uplus (3/4, 1] = \ldots$$

In Example 4.1.4 the subcomplexes of K may touch (that is, we can have $\overline{K_w} \cap \overline{K_{w'}} \neq \emptyset$ and $K_w \cap K_{w'} = \emptyset$), but in Example 4.1.3 they cannot. We introduce some terminology to capture the difference. Define

$$\mathcal{O} = \{x \in K \,;\, \exists\, n \in \mathbb{N}, w, w' \in W_n, \text{ such that } w \neq w' \text{ and } x \in \overline{K_w} \cap \overline{K_{w'}}\}. \tag{4.1.4}$$

Definition 4.1.5 *If K satisfies [$\mathscr{K}1$]-[$\mathscr{K}3$], and $\mathcal{O} = \emptyset$, then we say K is completely segregated.*

Example 4.1.3 is completely segregated, whereas Example 4.1.4 is not. None of our main results will require that K is completely segregated, but it is a useful concept and we will make reference to it several times in the sequel.

Lemma 4.1.6 *If K is completely segregated then for all $w \in W_*$, K_w is both closed and open.*

PROOF: If $\mathcal{O} = \emptyset$ then for all $n \in \mathbb{N}$, $K = \biguplus_{w \in W_n} K_w$ implies that $K = \bigcup_{w \in W_n} \overline{K_w}$, but $\overline{K_w} \cap \overline{K_{w'}} = \emptyset$ for $w \neq w' \in W_n$. Hence $K_w = \overline{K_w}$ is closed. Hence, $K_w = K \setminus (\bigcup_{w' \in W_n \setminus \{w\}} K_{w'})$ is open. ■

For the remainder the thesis we assume that K is a segregated space. We allow a more general form for $\mathcal{U}$ than the uniform measure used in Section 1.2. To be precise, let $\mathcal{U}$ be a measure on K which satisfies the following two conditions.

[$\mathscr{K}4$] $\mathcal{U}(K_w) > 0$ for all $w \in W_*$.

[$\mathscr{K}5$] *There exists a constant $\alpha \in (0,1)$ such that for all $w \in W_*$,*

$$\mathcal{U}\left(\bigcup \{K_{wi}\,;\, i \in S, \overline{K_{wi}} \not\subseteq K_w\}\right) \leq \alpha\, \mathcal{U}(K_w). \tag{4.1.5}$$

Both these conditions are satisfied by Examples 4.1.3 and 4.1.4 with $\mathcal{U} = \lambda$.

Recall from Section 1.2 that the main parameter our of model is a sequence $(r_n)_{n \geq 0}$ in $[0, \infty)$. To avoid degeneracy we require that $r_n > 0$ for some $n \in \mathbb{N}$. Recall that $\mathcal{R}(dw)$ is the measure on the countable set W_* which is defined by the point-masses $\mathcal{R}(w) = r_{|w|}$ and that $\mathcal{U}_w$ is the conditional measure of $\mathcal{U}$ on K_w defined by $\mathcal{U}_w(A) = \frac{\mathcal{U}(A \cap K_w)}{\mathcal{U}(K_w)}$. As in Section 1.2, let M be a Poisson point process M with points $(t, w, p) \in (-\infty, \infty) \times W_* \times K$ and rate $dt \otimes \mathcal{R}(dw)\mathcal{U}_w(dk)$. We write $(\Omega, \mathcal{F}, \mathbb{P})$ for the probability space on which M is defined. For (measurable) $I \subseteq \mathbb{R}$ and $V \subseteq W_*$ define

$$\begin{aligned} M_I &= \{(t, w, x) \in M\,;\, t \in I\} \\ M_{I \times V} &= \{(t, w, x) \in M\,;\, t \in I, w \in V\}. \end{aligned} \tag{4.1.6}$$

The condition [$\mathscr{K}4$] is required for our definition of M, in particular so as the conditional measure $\mathcal{U}_w$ is well defined. Loosely speaking, [$\mathscr{K}5$] is satisfied when the points sampled from $\mathcal{U}$ do not concentrate near a non-closed (in the topological sense) edge of any complex. By Lemma 4.1.6, [$\mathscr{K}5$] is trivially satisfied if K is completely segregated. Condition [$\mathscr{K}5$] prevents pathological examples of the parent sampling mechanism. The need for [$\mathscr{K}5$] will become apparent in the next section.

In the next section it will be to our advantage to have some almost sure properties of M as 'sure' properties of M. Define,

$$\mathcal{A}_{n,k} = \left\{ M_{[-k,k]\times W_n} \text{ is finite} \right\}$$
$$\mathcal{B} = \left\{ \text{for all } u \in \mathbb{R}, M_{\{u\}} \text{ is at most a single point} \right\}$$

By standard properties of Poisson point processes, $\mathbb{P}[\mathcal{A}_{n,k}] = 1$. For any two $w, w' \in W_*$, the probability of the processes $M_{\mathbb{R}\times\{w\}}$ and $M_{\mathbb{R}\times\{w'\}}$ causing a reproduction event at a common time is 0. Since W_* is countable, in fact also $\mathbb{P}[\mathcal{B}] = 1$. Hence, with slight abuse of notation, we simply redefine M so as

$$\Omega = \mathcal{B} \cap (\cap_{n,k} \mathcal{A}_{n,k}) . \tag{4.1.7}$$

We continue to denote the probability space on which M is defined by $(\Omega, \mathcal{F}, \mathbb{P})$.

4.2 Definition

In this section we give a mathematically precise definition of our model. We seek to make sense of the intuition which is conveyed by Definition 1.2.1 and Figure 1.1.

We begin with some discussion which motivates the formal definition. Let us briefly suppose that the K is completely segregated and that total event rate is finite; $\sum_n \mathcal{S}^n r_n < \infty$. Then $M_{[0,t]}$ is almost surely finite and we can represent the process over $[0, t]$ as in Figure 1.1. By following the arrows in Figure 1.1, we see precisely where we would like the flow X to map each $x \in K$ to, at time $t > 0$. We will shortly introduce the proper notation for doing this.

What is important to note in Figure 1.1 is which events affected the final position of the lineages[1]. Consider an event (s, w, y) in a complex K_w of level $|w| = n$ at time $s \in (0, t)$. The event had no effect on the final position (at time t) of any of the lineages if:

[1]For convenience, this information can be seen in Figure 4.1, overleaf.

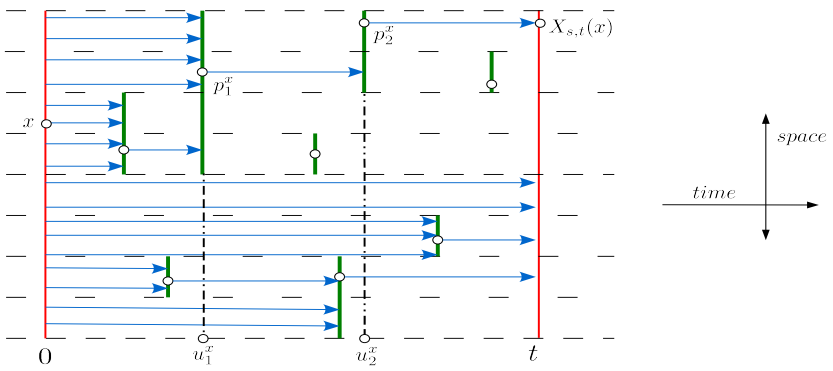


Figure 4.1: The sequences $(u_m^{x,0,t})$ and $(p_m^{x,0,t})$ corresponding to the point $x \in K$ are also shown (to illustrate Definition 4.2.1). In this example, $N_1^x = 1$, $N_2^x = 2$, and $N_3^x = \infty$.

- There was an event (s', w', y') such that $s < s' < t$ and $K_w \subseteq K_{w'}$.
- Or, the final event (s', w', y') such that $0 < s' < s$ and $K_w \subset K_{w'}$ had $y' \notin K_w$.

Hence, to work out where a point $x \in K$ should be mapped to (i.e. where, from time s, the individual at x has its ancestor at time $t > s$), we need only consider the following sequence of events.

First, we look for the final level 0 event during $(s, t]$ which affected the point x. If we find one, say (u_1, w_1, p_1), we then look for the final level 1 event which was after time u_1, and affected p_1. And so on. If at any point we don't find a level n event, we simply move up to the next level $n + 1$ and look there. Of course, if $\sum r_n < \infty$ then, at stage m of the induction, p_m might not have been affected by a single event over $(s_m, t]$, in which case we record this and stop the induction.

The formal embodiment of the preceding paragraph is as follows, and a graphical demonstration of the definition can be found in Figure 4.1.

Definition 4.2.1 *Fix (x, s, t) with $x \in K$ and $s < t$. Define $(u_0, w_0, p_0) = (s, \emptyset, x)$, and then for as long as $N_{m+1} < \infty$ define the (possibly finite) sequences*

$$N_1 = \inf\{n \geq 0\,;\, \exists (u, w, p) \in M_{(s,t]\times W_n} \text{ such that } p_0 \in K_w\}$$

$$E_1 = (u_1, w_1, p_1)$$

$$\text{where } u_1 = \max\{u \in (s, t]\,;\, (s, w, p) \in M_{(s,t]\times W_{N_1}} \text{ and } p_0 \in K_w\}.$$

$$N_{m+1} = \inf\{n > N_m \,;\, \exists (u,w,p) \in M_{[u_m,t]\times W_n} \text{ such that } p_m \in K_w\}$$

$$E_{m+1} = (u_{m+1}, w_{m+1}, p_{m+1})$$

$$\text{where } u_{m+1} = \max\{u \in [u_m, t] \,;\, (u,w,p) \in M_{[u_m,t]\times W_{N_{m+1}}} \text{ and } p_m \in K_w\}.$$

Remark 4.2.2 *By* (4.1.7)*, Definition 4.2.1 make sense for every $x \in K$ and $s < t$. In particular, knowing u_m uniquely specifies w_m and p_m.*

The (finite or infinite) sequence of events $(E_m) \subseteq M_{(s,t]}$ contains the only events which affected the final position of the lineage started from x over time $(s,t]$. At this point we make the key observation; for any $x \in K$, the sequence $(E_m) = (E_m)_{m\geq 1}$ is well defined *even* if we remove the condition that $\sum_n \mathcal{S}^n r_n$ be finite. In general, we have the following lemma.

We use the notation $s_n \uparrow s$ to mean that the sequence $(s_n) \subseteq \mathbb{R}$ is strictly monotone increasing, and $s_n \to s$.

Lemma 4.2.3 *If $\sum r_n < \infty$ then (E_m) is almost surely a finite sequence, $(E_m)_1^{\mathcal{M}}$, and $N_{\mathcal{M}+1} = \infty$. Further, in this case (E_m) is empty with positive probability.*

If $\sum r_n = \infty$, then, almost surely,

- *the sequence (E_m) is countably infinite, and $N_m \uparrow \infty$.*
- *$u_m \uparrow t$, $|w_m| \uparrow \infty$ and (p_m) converges.*

PROOF: If $\sum_n r_n < \infty$, note each point $x \in K$ is affected by reproduction events at rate $\sum_n r_n < \infty$. Thus the times in between the (u_m) are bounded below by exponential random variables with parameter $\sum_n r_n$.

Now suppose $\sum_n r_n = \infty$. To prove that (E_m) is infinite, note that for any $p \in K$

$$\{(u,w,p) \in \mathbb{R} \times W_* \times K \,;\, |w| \geq m, u \in (u_m, t), p \in K_w\}$$

has $dt \otimes \mathcal{P}(dw, dp)$ measure $(t - u_m) \times \sum_{n\geq m} r_n = \infty$, so the probability that p_m is hit by no events in (s_m, t) is zero.

To prove the remaining statements, suppose that for some $\epsilon > 0$, for all $m \in \mathbb{N}$, $u_m \leq t - \epsilon$. Then, by definition of (E_m), for all $m \geq \mathcal{M}$ there are no points in $M_{(t-\epsilon,t)}$, which is a contradiction. Hence $u_m \uparrow t$. Since $|w_m| = N_m$ and $N_{m+1} \geq 1 + N_m$, it is clear

that $|w_m| \to \infty$. Finally, note that K_{w_m} is a decreasing sequence of sets, and in fact $\sup\{m(y,z)\,;\, y,z \in K_{w_m}\} \leq L_{|w_m|} \to 0$. For all $k \geq m$, $p_k \in K_{w_m}$, and it follows from [$\mathscr{K}2$] that (p_m) is Cauchy, hence convergent, in K. ■

Notation 4.2.4 (Continuation of Definition 4.2.1) *The sequence (E_m) depends on x, s and t, and when we need this distinction (which will be most of the time) we write*

$$E_m^{x,s,t} = (u_m^{x,s,t}, w_m^{x,s,t}, p_m^{x,s,t}).$$

We write $K_{w_m^{x,s,t}} = K_{w_m}^{x,s,t}$.

We will use the sequence $(E_m^{x,s,t})$ to define $x \mapsto X_{s,t}(x)$ in (4.2.2), but there are some technicalities to work around before we can do this.

If K is completely segregated then (by Lemma 4.1.6) all complexes K_w of K are closed. Therefore, it makes intuitive sense that reproduction events occurring in complexes $K_{w'} \subseteq K_w$ cannot move particles in the flow from within K_w into $K \setminus K_w$.

If K is not completely segregated then it might be the case that an infinite sequence (u'_m, w'_m, p'_m) of events, with $K_{w'_m} \subseteq K_w$, could have $\lim p'_m \notin K_w$, because it could be that $\lim p'_m \in \overline{K}_w \setminus K_w$. In this case our construction would run into a serious problem; the flow property $X_{s,v} = X_{t,v} \circ X_{s,t}$ would fail. Thanks to [$\mathscr{K}5$], we are able to prove that $\lim p'_m$ is always in K_w, as the following result shows.

Define $\mathcal{C} \subseteq \Omega$ by

$$\mathcal{C} = \left\{ \lim_{m\to\infty} p_m^{x,s,t} \in K_{w_n}^{x,s,t} \text{ for all } n \in \mathbb{N} \text{ and } x,s,t \text{ such that } |E^{x,s,t}| = \infty \right\}. \tag{4.2.1}$$

Lemma 4.2.5 $\mathbb{P}[\mathcal{C}] = 1$.

Because of Lemma 4.2.5, it is advantageous to only use $(E_m^{x,s,t})$ for $\omega \in \mathcal{C}$. The proof of Lemma 4.2.5, which includes an application of [$\mathscr{K}5$], is given in Section 4.4. Note that, by Lemma 4.1.6, if K is completely segregated then all the K_w are closed and Lemma 4.2.5 is trivially true.

Briefly suppressing dependence on (s,t), note that $(K_{w_m}^x)$ (for $\omega \in \mathcal{C}$) is a decreasing infinite sequence of sets. We have also that $p_m^x \in K_{w_m}^x$ and $\sup\{m(y,z)\,;\, y,z \in K_{w_m}^x\} \leq L_{|w_m^x|}$. If (p_m^x) is infinite then by completeness of K and [$\mathscr{K}2$], (p_m^x) is convergent.

So, for $\omega \in \mathcal{C}$, define

$$X_{s,t}(x) = \begin{cases} x & \text{if } N_1^{x,s,t} = \infty \\ p_{\mathcal{M}}^{x,s,t} & \text{if } N_{\mathcal{M}+1}^{x,s,t} = \infty, \text{ for } \mathcal{M} \in \mathbb{N} \\ \lim_{m\to\infty} p_m^{x,s,t} & \text{if } (E_m^{x,s,t}) \text{ is infinite.} \end{cases} \tag{4.2.2}$$

and for $\omega \notin \mathcal{C}$ simply set $X_{s,t} = \iota$, the identity function on K.

Let D_K denote the Euclidean metric on $K \subseteq [0,1]$. Let $\mathbb{M}$ be the metric space of functions mapping K into K, equipped with the metric $||f,g||_\infty = \sup\{D_K(f(x),g(x))\,;\, x,y \in K\}$. Let $D_{[0,\infty)}(K)$ denote the space of càdlàg paths from $[0,\infty) \to K$, equipped with the usual Skorokhod topology (see Ethier and Kurtz 1986).

Theorem 4.2.6 *For each $s < t$, $X_{s,t}$ is an $\mathbb{M}$-valued random variable.*

- *For all $s < t < v$, $X_{s,v} = X_{t,v} \circ X_{s,t}$.*
- *For all $s < t < u < v$, $X_{u,v}$ and $X_{s,t}$ are independent.*
- *For all $t_1 - s_1 = t_2 - s_2$, X_{s_1,t_1} and X_{s_2,t_2} are identically distributed.*

The formula $X_{s,u} = X_{t,u} \circ X_{s,t}$ is known as the flow property and shows that the population which our model describes has a consistent genealogical structure. Lemma 4.2.5 and Theorem 4.2.6 are proved in Section 4.4 and several measurability/continuity properties of the model which we need for later chapters are established in Section 4.5. Before doing all this we will use the next section to give a short description of X when viewed backwards in time, establishing a dual process to our model.

4.3 The Segregated Λ-Fleming-Viot process

The Segregated Λ-coalescent is dual to a (not the) Spatial Λ-Fleming-Viot process, which we refer to in this section as the Segregated Λ-Fleming-Viot process. Let $\mathcal{K}$ be a compact metric space, from which the genetic types of individuals in the Segregated Λ-Fleming Viot process will be drawn.

In the Segregated Λ-Fleming-Viot process each site $x \in K$ carries a single individual. Without going into details regarding a state space and topology, we represent the Λ-Fleming-Viot process in similar style to the SΛFV process, as a function $\xi_t : K \to \mathcal{K}$. The type of the individual at x at time t is given by $\xi_t(x)$.

From some initial state $\xi_0 : K \to \mathcal{K}$, we define

$$\xi_t(x) = \xi_0(X_{-t,0}(x)). \tag{4.3.1}$$

From Definition 1.2.1 we can represent (ξ_t) in terms of a Poisson point process as follows. Define $\widetilde{M} = \{(-t, w, p)\,;\, (t, w, p) \in M\}$ and note that $\widetilde{M}$ is a Poisson point process with distribution identical to M. Initially, the individual at x has type $\xi_0(x) \in \mathcal{K}$. Then:

- Whenever $(t, w, p) \in \widetilde{M}$, a parent is sampled uniformly from within K_w and its offspring instantaneously colonize K_w. That is, $\xi_t(x) = \xi_{t-}(p)$ for all $x \in K_w$.

- In between reproduction events, the type of the individual at each $x \in K$ remains constant.

It is clear from the description above that the Segregated Λ-Fleming-Viot process is very similar to the SΛFV process with uniform killing, which we described in Section 3.8. The difference is that we have forced reproduction events to take place in precisely the complexes of K rather than just in any ball of K.

One important difference is that we can allow reproduction events to happen arbitrarily quickly, because $(X_{s,t})$ is defined for an arbitrary sequence (r_n). Better still, thanks to the flow $(X_{s,t})$, we can track all the ancestral lineages of the process together, so we do not run into the issues described in Section 3.7. Equation (4.3.1) gives duality in a very strong sense between the Segregated Λ-coalescent and the Segregated Λ-Fleming-Viot processes; the ancestral lineages of the Segregated Λ-Fleming-Viot process are precisely the particle paths of the Segregated Λ-coalescent.

The segregation of K into complexes, coupled with the fact that reproduction events occur only in the complexes, means that the probabilistic structure of the Segregated Λ-Fleming-Viot process is significantly more tractable than that of the spatial Λ-Fleming-Viot process. Note in particular that, in our model, a large number of small reproduction events cannot cause a particle to 'creep' far away from its starting point. A further illustration of tractability is found in a connection to Galton-Watson processes in varying environments which is described in Chapter 6 (we also touched upon this connection in Section 1.2).

4.4 Existence of the flow

The final two sections of this chapter are concerned with proving Theorem 4.2.6 and developing some measurability/continuity properties of our model. To do so we must work with the inductive, particle-wise definition of our model which we gave in Section 4.2. This necessitates some quite technical arguments so the reader may wish to omit the proofs given in the remainder of this chapter and return later. That said, once we have established the results of this section and Section 4.5, the particle-wise definition is an asset which becomes very useful in later chapters.

In this section we establish the existence of our model, in that we prove Lemma 4.2.5 and then Theorem 4.2.6.

PROOF: [of Lemma 4.2.5] Let

$$H' = \{(x, s, t) \in K \times \mathbb{R}^2 \,;\, |E^{x,s,t}| = \infty\}$$

and define an equivalence relation on H' by

$$(x, s, t) \sim (x', s', t') \Leftrightarrow E_1^{x,s,t} = E_1^{x',s',t'}.$$

Denote the equivalence class of (x, s, t) under $\sim$ by $[x, s, t]$. Note that for all $(x', s', t') \in [x, s, t]$, $(E_m^{x,s,t}) = (E_m^{x',s',t'})$. We thus write $u_m^{x,s,t} = u_m^{[x,s,t]}$, and similarly for w and p.

Note also that if $x' \in K_{w_1}^{x,s,t}$, $s' < u_1^{x,s,t}$ and $u_1^{x,s,t} < t' < t$ then $(x', s', t') \in [x, s, t]$. Let $(\hat{x}_k)_{k\in\mathbb{N}}$ be a deterministic countable subset of K, such that for all $w \in W_*$ there is $k \in \mathbb{N}$ such that $x_k \in K_w$. Let $(\hat{s}_k, \hat{t}_k)$ be a countably dense deterministic subset of $\{(s, t) \in \mathbb{R}^2 \,;\, s < t\}$. With a slight abuse of notation, enumerate their Cartesian product as

$$(\hat{x}_k, \hat{s}_k, \hat{t}_k)_{k\in\mathbb{N}}.$$

Thus, for all equivalence classes $[x, s, t]$ there exists (random) $k \in \mathbb{N}$ such that $(\hat{x}_k, \hat{s}_t, \hat{t}_k) \in [x, s, t]$.

Hence,

$$\begin{aligned}
\{\exists x, s, t \in H', \exists n \in \mathbb{N},\ &\lim_{m\to\infty} p_m^{x,s,t} \notin K_{w_n}^{x,s,t}\} \\
&= \{\exists [\hat{x}_k, \hat{s}_k, \hat{t}_k] \subseteq H', \exists n \in \mathbb{N},\ \lim_{m\to\infty} p_m^{[x,s,t]} \notin K_{w_n}^{[x,s,t]}\} \\
&= \bigcup_{k\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} \{\lim_{m\to\infty} p_m^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]} \notin K_{w_n}^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}\}. \qquad (4.4.1)
\end{aligned}$$

By $[\mathscr{K}1]$, $\lim_m p_m^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]} \notin K_{w_n}^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}$ occurs only if, for all $m \in \mathbb{N}$, $p_m^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}$ is such that

$$\overline{K_{w_{m+1}}^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}} \setminus K_{w_n} \neq \emptyset.$$

By $[\mathscr{K}5]$, for all $k, n \in \mathbb{N}$, for each fixed m the probability of this occurring is bounded above by $\alpha < 1$. The points $p_m^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}$ are independent random variables, each sampled according to the conditional measure of $\mathcal{U}$ on $K_{w_m}^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}$. Hence, $\mathbb{P}\left[\lim_{m\to\infty} p_m^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]} \notin K_{w_n}^{[\hat{x}_k,\hat{s}_k,\hat{t}_k]}\right] = 0$, and it follows from (4.4.1) that

$$\mathbb{P}\left[\exists x, s, t \in H', \exists n \in \mathbb{N},\ \lim p_m^{x,s,t} \notin K_{w_n}^{x,s,t}\right] = 0.$$

Thus $\mathbb{P}[\mathcal{C}] = 1$. ■

PROOF: [of Theorem 4.2.6] We omit a formal proof of the fact that $X_{s,t}(\cdot)$ is an $\mathbb{M}$-valued random variable, since it is clear from (4.2.2) that in the definition of $X_{s,t}$ (for fixed s, t) we have only used the $\mathcal{F}$-measurable sets $\mathcal{A}_{\setminus,\|}$ and $\mathcal{B}$, countably many (deterministic) operations, and the random variable $\mathcal{M}$.

The remainder of the proof comes in three parts, which correspond to the bullet points in the statement of Theorem 4.2.6. The first part is a careful check of the flow property.

Remark 4.4.1 *Although the flow property of X may seem intuitively obvious from Figure 1.1, to prove Theorem 4.2.6 we must work with the definition of X given in Section 4.2.*

PART 1. Let $s < t < v$ and fix $x \in K$. If $\omega \notin \mathcal{C}$ then we trivially have $X_{s,v} = X_{t,v} \circ X_{s,t}$. So, suppose $\omega \in \mathcal{C}$, and for convenience write $y = X_{s,t}(x)$. When necessary we will emphasise the dependence with $y = y^{x,s,t}$. We divide into three cases.

If $N_1^{x,s,t} = N_1^{y,t,v} = \infty$, then for all $x \in K$, $X_{s,t}(x) = x = y$ and $X_{t,v}(y) = y$. Since $x = y$, $N_1^{x,s,v} = \infty$ and $X_{s,t}(x) = X_{t,v}(x) = X_{s,v}(x)$, so $X_{t,v}(X_{s,t}(x)) = X_{s,v}(x)$.

If $N_1^{x,s,t} = \infty$ and $N_1^{x,t,v} < \infty$, then $X_{s,t}(x) = x = y$ and hence we must have $u_1^{x,s,t} \geq v$. Hence $(E_m^{x,t,v}) = (E_m^{x,s,v})$ and thus $X_{t,v}(X_{s,t}(x)) = X_{t,v}(x) = X_{s,t}(x)$.

If $N_1^{x,s,t} < \infty$, then we have $N_1^{x,s,v} < \infty$. Let

$$\mathcal{C}^{s,t,v} = \{x \in K\,;\, \exists m, u_m^{x,s,v} \geq t\}.$$

If $x \notin \mathcal{C}^{s,t,v}$ then $u_m^{x,s,v} < t$ for all m, so from the definitions we have $(E_m^{x,s,t}) = (E_m^{x,s,v})$. Hence $X_{t,v}(x) = X_{s,t}(x) = y$. Suppose it was the case that $(u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v}) \in (E^{y,t,v})$.

Note $y \in K^{x,s,v}_{w_1}$ so we must have $(u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v}) \in (E_m^{x,s,v})$, which is a contradiction since $u_1^{y,t,v} \geq t$. Hence $(E_m^{y,t,v})$ is empty, and $X_{t,v} = \iota$. Thus, $X_{t,v}(X_{s,t}(x)) = X_{s,v}(x)$.

If $x \in \mathcal{C}^{s,t,v}$, let

$$\mathcal{M} = \max\{m\,;\, u_m^{x,s,v} < t\}$$

(which is well defined since $((u_m^{x,s,t})$ is strictly increasing), and from the definitions note that $(E_m^{x,s,t})_1^{\mathcal{M}} = (E_m^{x,s,v})_1^{\mathcal{M}}$.

By definition of $\mathcal{M}$ we have $u^{x,s,v}_{\mathcal{M}+1} \geq t$ and, since $p^{x,s,v}_{\mathcal{M}} = p^{x,s,t}_{\mathcal{M}}$, it holds that $N^{x,s,v}_{\mathcal{M}+1} \leq N^{x,s,t}_{\mathcal{M}+1}$. Hence $K^{x,s,t}_{w_{\mathcal{M}+1}} \subseteq K^{x,s,v}_{w_{\mathcal{M}+1}}$. By definition, $p^{x,s,t}_{\mathcal{M}} \in K^{x,s,t}_{w_{\mathcal{M}+1}}$ and, we have also that $(K^{x,s,t}_{w_m})$ is decreasing. We have already commented that $K^{x,s,t}_{w_{\mathcal{M}+1}} \subseteq K^{x,s,v}_{w_{\mathcal{M}+1}}$, so it follows from Lemma 4.2.5 that $y^{x,s,t} \in K^{x,s,v}_{w_{\mathcal{M}+1}}$.

Since both y and $p^{x,s,v}_{\mathcal{M}}$ are elements of $K^{x,s,v}_{w_{\mathcal{M}+1}}$, there is no $(u, w, p) \in (E_m^{y,t,v})$ such that $|w| < N^{x,s,v}_{\mathcal{M}+1}$ - such a (u, w, p) would also have featured in $(E_m^{x,s,v})$, which contradicts the definition of $\mathcal{M}$. Also, there are no $(u, w, p) \in (E_m^{y,t,v})$ such that $u > u^{x,s,v}_{\mathcal{M}+1}$ and $y \in K_w$ - such a (u, w, p) would feature in $(E_m^{x,s,v})$, which contradicts the definition of $u^{x,s,v}_{\mathcal{M}+1}$.

Combining the results of previous two sentences, $(u^{x,s,v}_{\mathcal{M}+1}, w^{x,s,v}_{\mathcal{M}+1}, p^{x,s,v}_{\mathcal{M}+1}) = (u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v})$. Hence $(E_m^{x,s,v})_{m \geq \mathcal{M}+1} = (E_k^{x,s,v})_{k \geq 1}$, which implies that $X_{t,v}(y) = X_{s,v}(x)$. This completes the third case.

Since x and ω were arbitrary, in all cases we have that for all $\omega \in \Omega$, $X_{s,v} = X_{t,v} \circ X_{s,t}$.

Part 2: Let $s_1 < t_1 \leq s_2 < t_s$. Since $M_{(s_1,t_1]}$ and $M_{(s_2,t_2]}$ are independent, and the construction of $X_{s,t}$ depended only on $M_{s,t}$, it follows immediately that X_{s_1,t_1} and X_{s_2,t_2} are independent.

Part 3: Let $s_1 < t_1$ and $s_2 < t_2$ with $t_1 - s_1 = t_2 - s_2$. Then $M_{(s_1,t_1]}$ and $M_{(s_2,t_2]} - (t_2 - t_1)$ are identical in distribution, from which it follows that X_{s_1,t_1} and X_{s_2,t_2} are also identical in distribution. ■

4.5 Measurability and continuity properties

Recall that our underlying Poisson point process M is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout this section we denote the dependence on $\omega \in \Omega$ of X by writing $X_{0,t}(\cdot)(\omega)$.

Due to our inductive, particle-wise definition of X in Section 4.2, it is not immediately clear that functions such as $(x, t) \mapsto X_{0,t}(x)$ are random variables (in some space), or are

even measurable. In fact, as the results of this chapter show, our model is quite a regular object.

Lemma 4.5.1 *K is separable.*

PROOF: By Lemma 4.1.2, each K_w is non-empty. For each K_w pick some point $x(w) \in K_w$ and define $\mathbb{D} = \{x(w)\,;\, w \in W^*\}$. Note that $\mathbb{D}$ is countable.

Let O be an open set of K. Since K is a metric space, for some $r > 0$ and $y \in K$, $B_r(y) \subseteq O$. For $y \in K$ and $n \in \mathbb{N}$ let $K_{(y,n)}$ be the unique complex K_w of K such that $|w| = n$ and $y \in K_w$. By [$\mathscr{K}2$], for some $n \in \mathbb{N}$ we have $L_n < r/2$, so that $K_{(y,n)} \subseteq B_r(y) \subseteq O$. By definition of $x(w)$ there is $w \in W_*$ such that $x(w) \in K_{(y,n)} \subseteq \mathcal{O}$.

Hence $\mathbb{D}$ is a countable dense subset of K. ■

Lemma 4.5.2 *The Borel σ-algebra on K is generated by the sets $(K_w)_{w\in W_*}$.*

PROOF: Let $\mathbb{B}(K)$ denote the Borel σ-algebra on K and recall that D_K denotes the metric on K. Recall that, for $y \in K$ and $n \in \mathbb{N}$ $K_{(y,n)}$ is the unique complex K_w of K such that $|w| = n$ and $y \in K_w$. Note that it follows by [$\mathscr{K}2$] that

$$\{y\} = \bigcap_{n\in\mathbb{N}} K_{w(y,n)}.$$

It is clear that $\sigma(K_w\,;\, w \in W_*) \subseteq \mathbb{B}(K)$, and we must prove the reverse inclusion.

Since K is a metric space, its topology is defined to be that which is generated by the open balls of K. By Lemma 4.5.1, K is separable, so it follows that any open subset of K can be written as a union of only countably many open balls of K. Hence $\mathbb{B}(K)$ is generated by the open balls of K. So the proof is complete if we can show that any open ball of K is contained in $\sigma(K_w\,;\, w \in W_*)$.

To this end, let $B_r(x) = \{y \in K\,;\, |y - x| < r\}$ be a fixed but arbitrary open ball in K. Note that

$$B_r(x) \supseteq \bigcup\{K_w\,;\, w \in W_*, K_w \subseteq B_r(x)\}$$

is tautologically true, and, since W_* is countable, the union on the right is countable. Now, suppose that $y \in B_r(x)$. Since $B_r(x)$ is open, for some $\epsilon > 0$ we have $B_y(\epsilon) \subseteq B_r(x)$. By [$\mathscr{K}2$], for some sufficiently large $n \in \mathbb{N}$ we have

$$K_{w(y,n)} \subseteq B_\epsilon(y) \subseteq B_r(x).$$

However, the above equation implies that $K_{w(y,n)} \in \{K_w\,;\, w \in W_*, K_w \subseteq B_r(x)\}$, so as $y \in \bigcup\{K_w\,;\, w \in W_*, K_w \subseteq B_r(x)\}$. Hence, in fact

$$B_r(x) = \bigcup\{K_w\,;\, w \in W_*, K_w \subseteq B_r(x)\}$$

and thus $B_r(x) \in \sigma(K_w\,;\, w \in W_*)$. The proof is complete. ■

The following lemma will be needed in Section 7.1.

Lemma 4.5.3 *For each $s < t$, $(x, \omega) \mapsto X_{s,t}(x)(\omega)$ is a measurable function from $K \times \Omega \to K$.*

PROOF: Since the definition of X was homogeneous in time, it suffices to consider the case $s = 0$. For $v \in W_*$ let

$$\mathcal{W}(v) = \{w' \in W_*\,;\, K_v \subseteq K_{w'}\}.$$

Fix $w \in W_*$. We note that

$$\begin{aligned}
&\{(x,\omega) \in K \times \Omega\,;\, X_{0,t}(x)(\omega) \in K_w\} \\
&\quad = \bigcup_{v \in W_*} K_v \times \{\omega \in \Omega\,;\, X_{0,t}(K_v) \subseteq K_w\} \\
&\quad = \bigcup_{v \in W_*} K_v \times \Big[\{\omega \in \Omega \setminus \mathcal{C}\,;\, K_u \subseteq K_v\} \\
&\qquad\qquad \uplus \Big(\mathcal{C} \cap \big\{\omega \in \Omega\,;\, \nexists (u', w', p') \in M_{(0,t]}, K_v \subseteq K_{w'}, p' \notin K_w, M_{[u',t]\times\mathcal{W}(v)} = \emptyset\big\}\Big)\Big]
\end{aligned}$$

Note that $\{\omega \in \Omega \setminus \mathcal{C}\,;\, K_u \subseteq K_v\}$ is either empty or equal to the measurable set $\Omega \setminus \mathcal{C}$. From the representation above, it follows that $\{(x, \omega) \in K \times \Omega\,;\, X_{0,t}(x)(\omega) \in K_w\}$ is an element of the product σ-algebra on $K \times \Omega$. Lemma 4.5.2 completes the proof. ■

Remark 4.5.4 *An apparent omission from this section is the result that for all $\omega \in \Omega$, $X_{s,t} : K \to K$ is a measurable function. This appears later (as Lemma 5.5.2) because the proof uses some definitions that won't appear until Chapter 5.*

Recall that $D_{[0,\infty)}(K)$ denotes the space of càdlàg paths mapping $[0, \infty) \to K$, equipped with the Skorokhod topology. By Theorem 3.5.6 of Ethier and Kurtz (1986), the completeness of K implies that $D_{[0,\infty)}(K)$ is also complete. For each $t \in [0, \infty)$ Let $\partial_t : D_{[0,\infty)}(K) \to K$ be the projection map $\partial_t(f) = f(t)$.

We equip the space $D_{[0,\infty)}(K)$ with its Borel σ-algebra, $\mathbb{B}(D_{[0,\infty)}(K))$. By Theorem 3.7.1 of Ethier and Kurtz (1986), if $\mathcal{T}$ is any dense subset of $[0,\infty)$,

$$\mathbb{B}(D_{[0,\infty)}(K)) = \sigma(\partial_t\,;\, t \in \mathcal{T}). \tag{4.5.1}$$

Lemma 4.5.5 *Let $\mathcal{T}$ be a dense subset of $[0,\infty)$. The Borel σ-algebra $\mathbb{B}(D_{[0,\infty)}(K))$ on $D_{[0,\infty)}(K)$ is generated by $\{C_{t,w}\,;\, t \in \mathcal{T}, w \in W_*\}$, where*

$$C_{t,w} = \{f \in D_{[0,\infty)}(K)\,;\, f(t) \in K_w\}.$$

PROOF: By Lemma 4.5.2,

$$\begin{aligned}\sigma(\partial_t) &= \sigma(\partial_t^{-1}(K_w)\,;\, w \in W_*)\\ &= \sigma(C_{t,w}\,;\, w \in W_*).\end{aligned}$$

From (4.5.1) we have that

$$\mathbb{B}(D_{[0,\infty)}(K)) = \sigma\big(\sigma(\partial_t)\,;\, t \in \mathcal{T}\big)$$

and the result follows. ■

Lemma 4.5.6 *For all $\omega \in \Omega$ and $x \in K$, $t \mapsto X_{0,t}(x)(\omega)$ is a càdlàg function from $[0,\infty) \to K$.*

PROOF: The method of the following proof is to approximate $t \mapsto X_{0,t}(x)$ with càdlàg functions, such that the property of being càdlàg is preserved in the limit. To avoid cumbersome notation we will give an outline of the argument.

For $\omega \notin \mathcal{C}$, the result is trivial. For $\omega \in \mathcal{C}$, for all $n \in \mathbb{N}$, define

$$N_m^{x,0,t,(n)} = \begin{cases} \text{undefined} & \text{if } N_m^{x,0,t} \text{ is not defined} \\ N_m^{x,0,t} & \text{if } m < n \text{ and } E_m^{x,0,t} \text{ is defined} \\ \infty & \text{if } m = n \text{ and } E_m^{x,0,t} \text{ is defined} \\ \text{undefined} & \text{if } m > n \text{ and } E_m^{x,0,t} \text{ is defined} \end{cases}$$

and, correspondingly,

$$X_{s,t}^{(n)}(x) = \begin{cases} x & \text{if } N_1^{x,s,t,(n)} = \infty \\ p_M^{x,s,t} & \text{if } n \geq \mathcal{M} \text{ and } N_{\mathcal{M}+1}^{x,s,t,,(n)} = \infty \\ p_n^{x,s,t} & \text{if } n < \mathcal{M} \text{ and } N_{\mathcal{M}+1}^{x,s,t,,(n)} = \infty \\ p_n^{x,s,t} & \text{if } (E_m^{x,s,t}) \text{ is infinite.} \end{cases} \tag{4.5.2}$$

The point of this definition is that we mimic the definition of $X_{0,t}$, but we ignore all reproduction events with level n and above. It is immediate from [$\mathscr{K}$2] and the definition of $(E_m^{x,0,t})$ that for all $t > 0$,

$$\sup\left\{D_K\left(X_{0,t}(x), X_{0,t}^{(n)}(x)\right)\,;\, x \in K\right\} \leq L_n.$$

Since only finitely many reproduction events of level $\leq n$ occur during $[0,t]$, it is easily seen that $t \mapsto X_{0,t}^{(n)}(x)$ is càdlàg. Since right/left limits/continuity are preserved by uniform approximation, it follows that $t \mapsto X_{0,t}(x)$ is càdlàg. ■

Lemma 4.5.7 *For each $x \in K$, the function $\omega \mapsto X_{0,\cdot}(x)(\omega)$ is a measurable function from $\Omega \to D_{[0,\infty)}(K)$.*

PROOF: By Lemma 4.5.6 we have that $t \mapsto X_{0,t}(x)$ is càdlàg. Let $\mathcal{T}$ be a countable dense subset of $[0,\infty)$ and let $t \in \mathcal{T}$. We note that, similar to the proof of Lemma 4.5.3,

$$\begin{aligned}
&\{\omega \in \Omega\,;\, X_{0,\cdot}(x)(\omega) \in C_{t,w}\} \\
&\quad = \{\omega \in \Omega\,;\, X_{0,t}(x) \in K_w\} \\
&\quad = \{\omega \in \Omega \setminus \mathcal{C}\,;\, x \in K_w\} \\
&\qquad \uplus \left(\mathcal{C} \cap \left\{\omega \in \Omega\,;\, \nexists (u', w', p') \in M_{(0,t]}, x \in K_{w'}, p' \notin K_w, M_{[u',t]\times\mathcal{W}(w)} = \emptyset\right\}\right)
\end{aligned}$$

where $\mathcal{W}(v)$ is as defined in the proof of Lemma 4.5.3. Hence, for all $t \in \mathcal{T}$,

$$\{\omega \in \Omega\,;\, X_{0,\cdot}(x)(\omega) \in C_{t,w}\} \in \mathcal{F}.$$

By Lemma 4.5.5, $\omega \mapsto X_{0,\cdot}(x)(\omega)$ is measurable. ■

Remark 4.5.8 *There is no reason to expect continuity of $x \mapsto X_{0,t}(x)$, as Example 4.1.4 shows. In this example, there is positive probability of seeing no level 0 events, but seeing events in both $(1/2, 1]$ and $[0, 1/2]$. We thus see, with positive probability, a discontinuity of $x \mapsto X_{0,t}(x)$ at $x = 1/2$.*

If each K_w is open (which occurs, for example, by Lemma 4.1.6 if K is completely segregated) then

$$\mathbb{E}\left[D_K(X_{0,t}(x), X_{0,t}(y))\right] \to 0$$

as $y \to x$, where d_K denotes the metric on K. To see this, if $\sum r_n < \infty$, choose N such that, with high probability, no events of below level N hit x over $[0,t]$. Thus, with high

probability, defining $w \in W_N$ *by* $x \in K_w$ *and noting* K_w *is open,* $X_{0,t}$ *is locally about* x *equal to the identity, which is continuous. If* $\sum r_n = \infty$ *then (a.s.) there exists* $N \in \mathbb{N}$ *and* (s, w, p) *such that* $(s, w, p) \in M_{\{x\} \times W_N}$*. Hence, locally about* x*,* $X_{0,t}$ *maps to a single point, and is continuous.*

Chapter 5

Phases of the Segregated Λ-coalescent

In this chapter we classify the phases of the Segregated Λ-coalescent in an analogous style to our classification of the phases of the Λ-coalescent in Section 2.3. As part of this classification, we identify the existence of a critical phase that contains surprising behaviour which is not seen within the phases of the Λ-coalescent. We are able to give simple conditions involving (r_n) and $\mathcal{S}$ which determine when five distinct phases of our model occur.

By Theorem 4.2.6 the flow X is time homogeneous. From now on we will be concerned only with $(X_{0,t})_{t>0}$.

5.1 Dust and blocks

We think of the partition of K into singletons as the initial state of a coalescent, and the flow $t \mapsto X_{0,t}(x)$ as the coalescent mechanism which coagulates the blocks. That is, $x, y \in K$ are in the same block at time t if $X_{0,t}(x) = X_{0,t}(y)$. Recall from 5.1.1 that

$$\mathcal{E}_w = \inf\{t > 0\,;\, \exists u \in W_*, p \in K \text{ such that } K_w \subseteq K_u \text{ and } (t, u, p) \in M\}. \tag{5.1.1}$$

is known as the exponential clock associated to K_w.

Definition 5.1.1 *Let* $\mathcal{D}_t = \{x \in K\,;\ \text{for all } w \in W_* \text{ with } x \in K_w,\ \mathcal{E}_w > t\}$.

Note that $X_{0,t}(x) = x$ for all $x \in \mathcal{D}_t$. Define an equivalence relation on $K \setminus \mathcal{D}_t$ as follows: $x \overset{t}{\sim} y$ if and only if the particles which, at time 0 are at x and y respectively, are mapped to the same point of K at time t by the flow defined in Definition 1.2.1 (i.e. if $X_{0,t}(x) = X_{0,t}(y)$).

Definition 5.1.2 *Let* A_t *be the set of equivalence classes of* $\overset{t}{\sim}$.

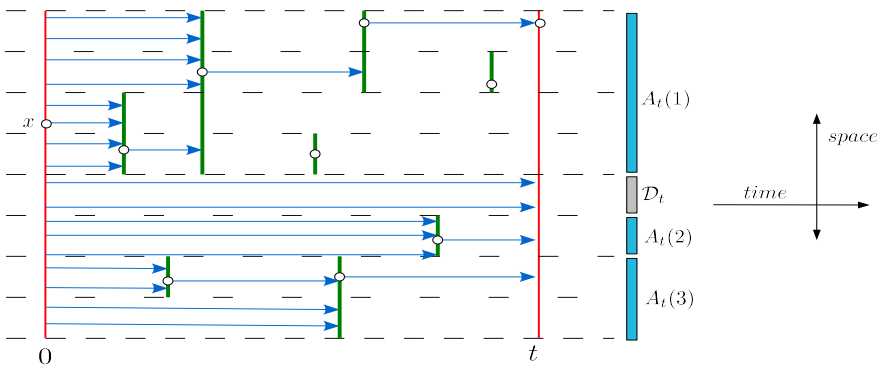


Figure 5.1: The division of K into dust and non-singleton blocks over time $[0, t]$ is shown. The dust is labelled $\mathcal{D}_t$, while the non-singleton blocks (three, in this realisation) are labelled $A_t = \{A_t(1), A_t(2), A_t(3)\}$.

Thus A_t is a partition of $K \setminus \mathcal{D}_t$. We call $\mathcal{D}_t$ the *dust* which is present at time t, and for each $x \in \mathcal{D}_t$ we say $\{x\}$ is a *singleton* (or singleton block) at time t. Each element of A_t is a *non-singleton* (or non-singleton block), and A_t is the set of non-singleton blocks. The division of K into dust and non-singleton blocks, in the same realisation as Figure 1.1, is shown in Figure 5.1. Note that

$$K = \mathcal{D}_t \uplus \left(\biguplus A_t \right). \tag{5.1.2}$$

That is, every point $x \in K$ is either part of the dust or is contained in some non-singleton block.

Loosely speaking, phase transitions in the Segregated Λ-coalescent occur as we increase the reproduction rate from a sequence where $r_n \downarrow 0$, through to a sequence where $r_n \uparrow \infty$. A phase transition means a change in qualitative behaviour of the singleton or non-singleton blocks. They may change behaviour together, or one behaviour may change while the other does not.

To be precise, we classify our phases according to whether the total number $|A_t|$ of non-trivial blocks is finite or countably infinite, and whether $\mathcal{D}_t$ is empty, non-empty and null, or takes up a positive fraction of the population. When considering the measure of $\mathcal{D}_t \subseteq K$, we mean with respect to the measure λ on K.

Note that A_t is at most countable, since at most countably many reproduction events happen during $(0, t]$. When $\mathcal{D}_t$ takes up a positive fraction of the population we say $\mathcal{D}_t$ is positive.

	Segregated Λ-coalescent	Λ-Coalescent
no dust	$\mathcal{D}_t = \emptyset$	$N^s = 0$
positive dust	$\lambda(\mathcal{D}_t) > 0$	$\mathscr{D}(\Pi_t) > 0$ (or $N^s = \infty$)
non-empty null dust	$\mathcal{D}_t \neq \emptyset$, $\lambda(\mathcal{D}_t) = 0$	(does not occur)

	Segregated Λ-coalescent	Λ-Coalescent
finitely many non-singletons	$\lvert A_t\rvert < \infty$	$N_t^a < \infty$
infinitely many non-singletons	$\lvert A_t\rvert = \infty$	$N_s^a = \infty$

Table 5.1: Comparison of the notation for the dust/blocks of the Λ-coalescent and Segregated Λ-coalescent.

Recall that in Chapter 2 we noted a similar decomposition for the particles in the Λ-coalescent; blocks of the Λ-coalescent were either singletons or contained a non-trivial proportion of the total population. The relationship between the notation used for the Λ-coalescent in Section 2.3 and the notation defined above can be seen in Table 5.1.

In the next section we discuss which combinations of the behaviours described in Table 5.1 actually occur.

5.2 The phase classification

The Segregated Λ-coalescent has five phases which we define here by means of the following table.

$\mathcal{D}_t$ is:	empty	positive	positive	non-empty, null
A_t is:	finite	finite	infinite	infinite
Lower subcritical	X	X		
Upper subcritical	X		X	
Semicritical	X			X
Critical	X			X (before t_0)
Supercritical	X			

Table 5.2: The phases of our model, and the behaviours occurring in each phase. A crossed box means that, in the corresponding phase, the corresponding behaviour occurs with positive probability. For any phase, the crossed boxes detail all the behaviour that occurs with positive probability.

In the critical phase the Segregated Λ-coalescent has the surprising property that for some *deterministic* $t_0 \in (0, \infty)$,

$$\mathbb{P}\left[\mathcal{D}_t \text{ is null but non-empty}, A_t \text{ is countably infinite}\right] > 0 \text{ for } t < t_0$$

but

$$\mathbb{P}\left[\mathcal{D}_t \text{ is empty}, A_t \text{ is finite}\right] = 1 \text{ for } t > t_0.$$

The time $t_0 \in (0, \infty)$ at which this transition occurs is called the *critical time*.

The critical phase of the model contains (but is not limited to) the case $r_n = c \in (0, \infty)$, where the GWVEs are in fact classical Galton-Watson processes. We will discuss the critical phase of the model in Chapter 8.

The classification of the phases of X is as follows.

Theorem 5.2.1 *Dependent only upon $\mathcal{S}$ and (r_n), our model is in precisely one of the five phases. In fact, X is*

- *lower subcritical if and only if $\sum_n \mathcal{S}^n r_n < \infty$.*
- *upper subcritical if and only if $\sum_n \mathcal{S}^n r_n = \infty$ and $\sum r_n < \infty$.*
- *semicritical if and only if $\sum r_n = \infty$ and $\limsup_n \frac{1}{n} \sum_1^n r_j = 0$.*
- *critical if and only if $\limsup_n \frac{1}{n} \sum_1^n r_j \in (0, \infty)$*
- *supercritical if and only if $\limsup_n \frac{1}{n} \sum_1^n r_j = \infty$.*

The proof of Theorem 5.2.1 will be given in Chapter 7.

Remark 5.2.2 *The sequence (r_n) controls the rate at which reproduction events occur, and the measure $\mathcal{U}$ controls the location of the parent points. However, by Definition 1.2.1, the parent sampling has no effect on the rate at which reproduction events occur since for all $w \in W_*$, $\mathcal{U}_w(K) = 1$. Thus $\mathcal{U}$ plays no part in determining the phase of our model; it does not appear in Theorem 5.2.1.*

Regardless of the phase, there is some random time $\tau < \infty$ which is (by definition) the first time at which a finite number of reproduction events have covered K. After this time our model assumes the behaviour designated by the leftmost column of Table 5.2. However, outside of the critical and supercritical phases τ is distributed over $(0, \infty)$. In the critical phase the distribution of τ has compact support $[0, t_0]$, whereas in the supercritical phase $\tau = 0$ almost surely.

As the phase of our model changes, the behaviour of the dust is as expected, in that increasing the intensity of reproduction events reduces the fraction of dust. The lack of monotonicity in the behaviour of the non-trivial blocks is explained as follows. In the lower subcritical phase there are simply not enough events to make anything more than finitely

many non-trivial blocks. Then, as the rate increases, there is an intermediate period where we see a countably infinity of non-trivial blocks. Eventually there are so many reproduction events that they frequently overlap, and we need (a.s.) only finitely many of them to cover K.

In Section 1.2 we commented that when reproduction events are occurring at a high rate, it is common for a larger reproduction event to overwrite the effect of some of the preceding smaller ones. This is borne out by appearance of the lim sup in the formula

$$\limsup_n \frac{1}{n} \sum_1^n r_j. \tag{5.2.1}$$

From Theorem 5.2.1 we see that, when $\sum r_n = \infty$, only the n-level reproduction events for which r_n is large enough to contribute to the lim sup take part in determining the phase.

Our proof of Theorem 5.2.1 is outlined in Section 5.4. As we will see, the quantity $\limsup_n \frac{1}{n} \sum_1^n r_j$ characterises the behaviour of the associated GWVEs.

5.3 Comparison to the phases of the Λ-coalescent

Recall that, if $\Lambda(\{1\}) > 0$, then the effect on the Λ-coalescent of the atom at 1 is as follows: Independently of all other mergers (and at rate $\Lambda(\{1\})$) the Λ-coalescent sees mergers which coagulate the whole population into a single block. Thus, the atom at 1 serves only to obfuscate the behaviour of the Λ-coalescent and in Chapter 2 we opted to simply remove it.

In the Segregated Λ-coalescent, at first glance it might seem that an equivalent operation would be to set $r_0 = 0$. However, for any $n \in \mathbb{N}$ with $r_n > 0$, there is positive probability of a finite number of level n reproduction events covering K before time t, and causing $|X_{0,t}(K)| \leq \mathcal{S}^n < \infty$. The result is that it is not possible to simply remove the possibility that a finite number of reproduction events caused the whole population to coagulate into a finite number of blocks. Thus, the first difference between the phase behaviour of the Λ-coalescent and the Segregated Λ-coalescent is the following.

1. Playing the role of the cases where $\Lambda(\{1\}) > 0$, we have the (always positive) probability of having only finitely many non-trivial blocks and no dust.

Having made this observation, we can now draw a diagram similar to Figure 2.1 of the phase transitions observed in the Λ-coalescent.

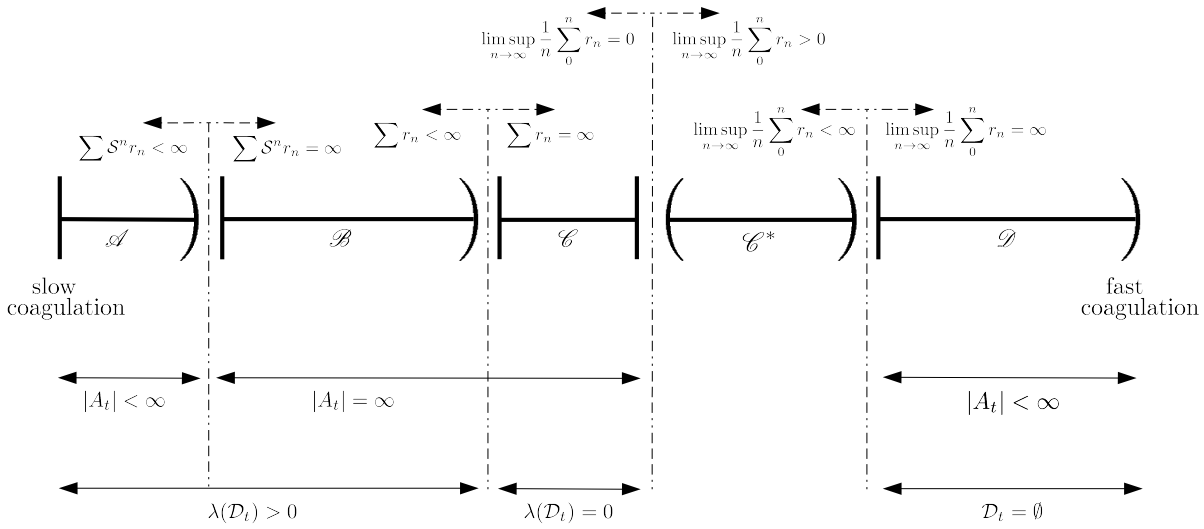


Figure 5.2: The phases of the Segregated Λ-coalescent are shown, with the understanding that (in addition, where necessary) there is always positive probability of having empty dust and only finitely many blocks. The phases are labelled $\mathscr{A} - \mathscr{D}$ to match their equivalents from Figure 2.1, with the critical phase labelled $\mathscr{C}^*$.

Through Table 5.1, the phases labelled $\mathscr{A}$-$\mathscr{D}$ in Figure 5.2 correspond to their equivalents in Figure 2.1, which are themselves also labelled $\mathscr{A}$-$\mathscr{D}$. Figure 5.2 also makes clear two important difference between the phases of the Λ-coalescent and our own model.

2. There is no possibility in our model of having a countable infinity of non-trivial blocks and empty dust. Instead, in the phase labelled $\mathscr{C}$, the (semicritical) Segregated Λ-coalescent sees a countable infinity of non-singleton blocks and non-empty null dust.

3. The critical phase of the Segregated Λ-coalescent, labelled $\mathscr{C}^*$, appears in between the semi- and supercritical phases. This phase has no equivalent from the behaviour of the Λ-coalescent.

Recall the spatial version of the Λ-coalescent from Angel et al. (2010) which was described in Section 2.4. Angel et al. showed that their coalescent model did not come down from infinity whereas, in contrast, the Segregated Λ-coalescent can come down from infinity.

Our model is very different to that of Angel et al. (2010) and we are now able to outline some big differences between the two. The geographical space in our model is a spatial continuum (instead of a discrete graph) and we permit only a single block to occupy each spatial location at any one time. The blocks in our model do not move in space, except

during mergers. Further, when a merger occurs in our model it involves all the blocks from within a non-null proportion of the geographical space. Thus, the multiple mergers in the two models are 'multiple' for quite different reasons, and it is natural to expect different behaviour.

5.4 Outline of the proof of the phase classification

In this section we give an outline of the proof of Theorem 5.2.1. The proof itself will be spread over Chapters 6 and 7.

Our proof of Theorem 5.2.1 comes via several lines of enquiry. Firstly Fubini's theorem produces some useful information, through much the same method as we used in Lemma 3.7.1. Secondly, the spatial structure of K eliminates some of the potential possibilities. These two approaches are (almost) enough to separate out the lower and upper subcritical phases, using the quantities $\sum_n \mathcal{S}^n r_n$ and $\sum_n r_n$.

A more powerful tool, namely the connection to GWVEs, is required to identify the semicritical, critical and supercritical phases. For example, Lemma 6.4.3 says that (a.s.)

$$\mathcal{D}_t = \bigcap_{n=0}^{\infty} \bigcup_{w \in \mathscr{B}_n^t} K_w. \tag{5.4.1}$$

Our definition of $\mathscr{B}_n^t$, which is that

$$\mathscr{B}_n^t = \left\{ w \in W_n \,;\ \text{for all } j = 0, 1, \ldots, n,\ \mathcal{E}_{w_1 \ldots w_n} > t \right\}$$

already suggests that the relationship (5.4.1) holds. To say that $x \in \bigcap_n \bigcup_{w \in \mathscr{B}_n^t} K_w$ is, by definition of $\mathscr{B}^t$, precisely the statement that $K_v \ni x$ implies $\mathcal{E}_v > t$. It is clear from (5.4.1) that the behaviour of $B_n^t = |\mathscr{B}_n^t|$ as $n \to \infty$ is closely connected to the behaviour of $\mathcal{D}_t$.

Remark 5.4.1 *Formulas similar to* (5.4.1) *can be found in the random fractals literature at least as far back as Falconer (1986), Mauldin and Williams (1986) and Graf (1987) (although these authors did not use branching processes explicitly). Such formulas provide what is now a well known connection between various classes of random fractals and branching processes.*

In the random fractals literature the complexes K_w are generally assumed to be closed, from which the completeness of K and [$\mathscr{K}2$] implies that $\cap_n K_{w(n)}$ is equal to a single point for each decreasing sequence $(K_{w(n)})_{n \in \mathbb{N}}$. For us this is not the case unless K is completely

segregated (recall Lemma 4.1.6). We are rescued by $\mathcal{U}$ and [$\mathcal{K}$5], which force enough of the $\cap_n K_{w(n)}$ to be non-empty that the relationship (5.4.1) *is still useful to us. See Remark 7.3.2 and the proof of Lemma 7.3.1.*

A GWVE (B_n) is said to be *degenerate* if $\mathbb{P}[\exists n \in \mathbb{N}, B_n = 0] = 1$. GWVEs may have multiple rates of growth and this complicates their behaviour considerably by comparison to the classical Galton-Watson process. Amongst others, D'Souza (1994) studies this phenomenon and gives an example of a GWVE with a countable infinity of growth rates.

A general degeneracy criterion for GWVEs seems not to be known. However, our offspring distributions are binomial with a fixed number of trails and in this case necessary and sufficient conditions were first given in Agresti (1975)[1]. A second set of conditions for degeneracy (in a slightly different setting) were obtained independently by Jirina (1976) and further conditions (in a slightly different setting again) were given by Lyons (1992).

Whichever set of conditions we use, there is still some work to be done in checking the conditions and thus determining which of the GWVEs we consider are degenerate. Jirina's conditions are slightly better suited to our situation than the others and we state them in Lemma 7.4.1.

The most important quantity involved in determining degeneracy of a GWVE is the limiting behaviour of its expectation. By analogy with the classical Galton-Watson process, we might hope that the limiting behaviour of $\mathbb{E}[B_n]$ was enough information to determine degeneracy, but for general GWVEs this is not the case. A large class of explicit examples which illustrate this issue, along with a wider discussion of conditions for degeneracy, can be found in Pemantle (1997).

Fortunately, Lemma 7.4.2 shows that for us the extra complications occur only in the critical phase, at the critical time. In particular, it turns out that degeneracy of B_n^t is equivalent to $\inf_n \mathbb{E}\left[B_n^t\right] = 0$, *except* in the critical phase when $t = t_0$. We will look at what happens in the critical phase at $t = t_0$ in Chapter 8.

A further piece of information, Lemma 7.3.5, comes out of the fact that, if $\sum r_n < \infty$, then B_n^t either becomes extinct or grows exponentially as $n \to \infty$. Using a result of Biggins and D'Souza (1992), we show that this fast growth implies that if $\mathcal{D} \neq \emptyset$ then $\mathcal{D}_t$ is non-null.

[1]To be precise, Agresti's conditions fail to cover our model when $r_n \to 0$, but this case can be dealt with via elementary means.

5.5 Some notation

In the notation of Section 4.2, we have that almost surely (i.e. for $\omega \in \mathcal{C}$),

$$\mathcal{D}_t = \{x \in K\,;\, N_1^{x,0,t} = \infty\}$$

since $N_1^{x,s,t} = \infty$ if and only if $(E_m^{x,s,t}) = \emptyset$. For $t > 0$ define

$$\mathscr{A}_t = \{X_{0,t}(x)\,;\, \exists w \in W_*, x \in K_w \in A_t\}. \tag{5.5.1}$$

It will be convenient for us to write $f^{-1}(\{x\}) = f^{-1}(x)$ for functions f.

Lemma 5.5.1 *The following statements hold:*

1. *For all $x \in \mathcal{D}_t$, $X_{0,t}^{-1}(x) = \{x\}$.*

2. *There is a bijective correspondence between $\mathscr{A}_t$ and A_t which is defined by*

$$y \mapsto X_{0,t}^{-1}(y),$$

for $y \in \mathscr{A}_t$. Further, for any $y \in \mathscr{A}_t$ and $x \in K$ such that $X_{0,t}(x) = y$,

$$X_{0,t}^{-1}(y) = K_{w_1}^{x,0,t}.$$

PROOF: Let $x \in \mathcal{D}_t$, then by definition $X_{0,t}(x) = x$. Suppose $z \in K \setminus \{x\}$ and $X_{0,t}(x) = z$. Then $(E_m^{z,0,t}) \neq \emptyset$. If $x \notin K_{w_1}^{z,0,t}$ then by Lemma 4.2.5 we would have $X_{0,t}(x) \neq X_{0,t}(z)$, so we must have $x \in K_{w_1}^{z,0,t}$. But this contradicts $x \in \mathcal{D}_t$. Hence in fact $X_{0,t}^{-1}(x) = x$, which proves our first statement.

If $\omega \notin \mathcal{C}$ then A_t and $\mathscr{A}_t$ are both empty, so our second statement holds trivially in this case. So, consider $\omega \in \mathcal{C}$. If $y \in \mathscr{A}_t$ then we must have $y = X_{0,t}(x)$ for some $x \in K \setminus \mathcal{D}_t$. By definition of A_t, $x \notin \mathcal{D}_t$. Let $B \in A_t$ be the unique block such that $x \in B$. Note that

$$B = X_{0,t}^{-1}(y)$$

by definition of A_t. By definition of X, we have also that $K_{w_1}^{x,0,t} \subseteq X_{0,t}^{-1}(y)$. If $z \in X_{0,t}^{-1}(y)$ and $z \notin K_{w_1}^{x,0,t}$, then Lemma 4.2.5 implies $X_{0,t}(z) \neq X_{0,t}(x) = y$, so in fact

$$K_{w_1}^{x,0,t} = X_{0,t}^{-1}(y)$$

for all $y \in \mathscr{A}_t$. By Lemma 4.2.5 and the definition of X, $X_{0,t}(z) = y$ for all $z \in K_{w_1}^{x,0,t}$, hence $K_{w_1}^{x,0,t} \subseteq B$. If $z \in B \setminus K_{w_1}^{x,0,t}$ then by definition of $K_{w_1}^{x,0,t}$ we would have $X_{0,t}(x) \neq X_{0,t}(x)$, but this contradicts the fact that $z \in B$. Hence in fact

$$B = K_{w_1}^{x,0,t}$$

and in particular, $K_{w_1}^{x,0,t} \in A_t$.

We now show that the function

$$y \overset{\phi_t}{\mapsto} X_{0,t}^{-1}(y) = K_{w_1}^{x,0,t}$$

is a bijection. By definition of A_t, for each $B' \in A_t$ there is a unique $y' \in K$ such that $X_{0,t}(B) = \{y'\}$. Hence also $X_{0,t}^{-1}(y') = B$ and we have that ϕ_t is surjective. If $y_1, y_2 \in \mathscr{A}_t$ are such that $X_{0,t}^{-1}(y_1) = X_{0,t}^{-1}(y_2)$ then for any $x \in X_{0,t}^{-1}(y_1)$ we have $y_1 = X_{0,t}(x) = y_2$ so as ϕ_t is also injective. This completes the proof. ■

Recall that the following result was promised in Remark 4.5.4.

Lemma 5.5.2 *Let $-\infty < s < t < \infty$. Then for all $\omega \in \Omega$, $X_{s,t} : K \to K$ is a measurable function.*

PROOF: Since our model is time homogeneous, it suffices to consider the case $s = 0$. If $\omega \notin \mathcal{C}$ then $X_{0,t}$ is the identity function, which is measurable.

If $\omega \in \mathcal{C}$, then $X_{0,t}|_{\mathcal{D}_t}$ is also the identity function, which is measurable. On $K \setminus \mathcal{D}_t$, $X_{0,t}$ has a countable range, which by Lemma 5.5.1 is precisely the set $\mathscr{A}_t$. Also by Lemma 5.5.1, if $y \in \mathscr{A}_t$ then $X_{0,t}^{-1}(y) = K_w$ for some $w \in W_*$. By definition of a segregated space, every K_w is measurable. Thus $K \setminus \mathcal{D}_t$ is a countable union of measurable sets, and is itself measurable. It follows immediately that $\mathcal{D}_t$ is measurable.

Thus the restriction $X_{0,t}$ is measurable on both $\mathcal{D}_t$ and $K \setminus \mathcal{D}_t$, both of which are measurable sets. It follows that $X_{0,t}$ is measurable. ■

We now set up some notation to describe Table 5.2. We write $\mathbb{P}\langle A \divideontimes B \rangle = 1$ to mean that $\mathbb{P}[A \cup B] = 1$ and both $\mathbb{P}[A] > 0$, $\mathbb{P}[B] > 0$. Define

$$\mathscr{P}_1^t = \{\mathcal{D}_t = \emptyset, A_t \text{ is finite}\}$$
$$\mathscr{P}_2^t = \{\lambda_t(\mathcal{D}_t) > 0, A_t \text{ is finite}\}$$

$$\mathscr{P}_3^t = \{\lambda_t(\mathcal{D}_t) > 0, A_t \text{ is countably infinite}\}$$

$$\mathscr{P}_4^t = \{\lambda_t(\mathcal{D}_t) = 0, \mathcal{D}_t \neq \emptyset, \text{ and } A_t \text{ is countably infinite}\}.$$

Definition 5.5.3 *We say that X is*

- *Lower subcritical if for all $t > 0$, $\mathbb{P}\big\langle \mathscr{P}_1^t \circledast \mathscr{P}_2^t \big\rangle = 1$.*
- *Upper subcritical if for all $t > 0$, $\mathbb{P}\big\langle \mathscr{P}_1^t \circledast \mathscr{P}_3^t \big\rangle = 1$.*
- *Semicritical if for all $t > 0$, $\mathbb{P}\big\langle \mathscr{P}_1^t \circledast \mathscr{P}_4^t \big\rangle = 1$.*
- *Critical if there exists $t_0 > 0$ such that:*
  - *For all $t \in (0, t_0)$, $\mathbb{P}\big\langle \mathscr{P}_1^t \circledast \mathscr{P}_4^t \big\rangle = 1$.*
  - *For all $t \in (t_0, \infty)$, $\mathbb{P}\left[\mathscr{P}_1^t\right] = 1$.*
- *Supercritical if $\forall t > 0$, $\mathbb{P}\left[\mathscr{P}_1^t\right] = 1$.*

Chapter 6

Connections to GWVEs

In this chapter we develop the connection between the Segregated Λ-coalescent and GWVEs which was touched upon in Section 1.2. To be precise, we formulate $\mathcal{D}_t$ and A_t in terms of a family of GWVEs associated to our model.

6.1 Galton-Watson processes in varying environments

Recall that a GWVE $(B_n)_{n\geq 0}$ is a classical Galton-Watson process with the modification that the offspring distribution of an individual may depend on its generation number.

To be precise, let $\mathcal{G}^{(0)} \in \mathbb{N}_0$ and let $(\mathcal{G}^{(n)})_{n\in\mathbb{N}}$ be a sequence of $\mathbb{N}_0$-valued random variables, called offspring distributions. Let $(\mathcal{G}_i^{(n)})_{i=1}^{\infty}$ be a sequence of independent copies of $\mathcal{G}^{(n)}$. We assume that for all $n \in \mathbb{N}_0$, $\mathbb{P}\left[\mathcal{G}^{(n)} = 0\right] < 1$.

The GWVE with offspring distribution $\mathcal{G}^{(n)}$ is defined by $B_0 = \mathcal{G}^{(0)}$ and the inductive relation

$$B_n = \sum_{i=1}^{B_{n-1}} \mathcal{G}_i^{(n-1)}.$$

The classical Galton-Watson process corresponds to the case where $\mathcal{G}^{(n)}$ has the same distribution, for all n.

The following result is well known and appears in Fearn (1972). Set $\mu_n = \mathbb{E}\left[\mathcal{G}^{(n)}\right]$ and $\sigma_n^2 = \mathrm{var}[\mathcal{G}^{(n)}]$.

Lemma 6.1.1 *It holds that*

$$E\left[B_n\right] = \prod_{l=0}^{n-1} \mu_n$$

$$\mathrm{var}\left[B_n\right] = E\left[B_n\right]^2 \sum_{l=0}^{n-1} \frac{\sigma_l^2}{\mu_l \prod_{k=0}^{l} \mu_k}$$

A consequence of this, which we will not require until Chapter 10, is

$$\begin{aligned}\operatorname{var}[B_n] &= \sum_{l=0}^{n-1}\left(\frac{1}{\mu_l}\prod_{k=0}^{n-1}\mu_k\right)\frac{\sigma_l^2\mu_l}{\mu_l}\frac{\prod_{k=0}^{n-1}\mu_k}{\prod_{k=0}^{l}\mu_k} \\ &\leq \left(\left(\min_{k=0,\ldots,n-1}\mu_k\right)^{-1}\prod_{k=0}^{n-1}\mu_k\right)\left(\max_{k=0,\ldots,n-1}\sigma_k^2\right)\sum_{l=0}^{n-1}\prod_{k=l+1}^{n-1}\mu_k. \qquad (6.1.1)\end{aligned}$$

Equation (6.1.1) allows us to control the variance of B_n using the expectations of $(B_m)_{m\leq n}$.

6.2 Notation for trees

Consider W_* as the set of nodes of an $\mathcal{S}$-ary tree, in the natural way. That is, for all $w \in W_*$ and $i \in S$ there is a directed edge (w, wi) connecting w to wi. Let E_* denote the set of edges of W_*, given by

$$E_* = \{(w, wi)\,;\, w \in W_*, i \in S\}. \qquad (6.2.1)$$

In the sequel, if $w \in W_*$ we will tend to write $w = w_1 \ldots w_n$ where $|w| = n$ and $w_j \in S$ are the letters of the word w. In this representation, we use the convention that $n \leq 0$ is equivalent to $w = \emptyset$.

A path on the $\mathcal{S}$-ary tree W_* is any subset of E_* which has the form

$$\{(v, vw_1), (vw_1, vw_1w_2), \ldots, (vw_1 \ldots w_{n-1}, vw\}$$

for some $v, w \in W_*$, where $w = w_1 \ldots w_n$. Such a path is said to connect v and vw, and the nodes on the path are said to be $v, vw_1, \ldots, vw$. We say a subset U of W_* is connected if, whenever $u, v \in U$ have a path connecting them, all the nodes on this path are also elements of U. We extend the definition of paths and nodes on paths to 'infinite paths' of the form $\{(v, w_1), (vw_1, vw_1w_2), \ldots\}$ in the obvious manner.

If U is a non-empty subset of W_*, then a point $w \in W_*$ is said to be a boundary point of U if there is some $u \in U$ and $i \in S$ such that $w = ui$ and $w \notin U$.

Lemma 6.2.1 *Let U be a finite connected subset of W_* and let $\overline{U}$ denote the set of boundary points of U. Then*

$$|\overline{U}| = (\mathcal{S} - 1)|U| + 1. \qquad (6.2.2)$$

PROOF: Note that (6.2.2) holds if U if a single point. Now, suppose that U is a finite non-empty connected subset of W_* and that $V = U \cup \{b\}$ where b is a boundary point of

U. Clearly $|V| = |U| + 1$. Since $b \in V$, V does not have b as a boundary point, but the $\mathcal{S}$ points which are boundary points of b are boundary points of V (and not of U). Hence V has $\mathcal{S} - 1$ more boundary points than U. Thus, if U satisfies (6.2.2) we have

$$\begin{aligned} |\overline{V}| &= |\overline{U}| + \mathcal{S} - 1 \\ &= (\mathcal{S} - 1)(|U| + 1) + 1 \\ &= (\mathcal{S} - 1)V + 1. \end{aligned}$$

Any non-empty connected subset of W_* can be constructed from a single point by repeated addition of boundary points; the result follows by induction. ■

6.3 GWVE trees

Definition 6.3.1 *Let v be some point in W_*. Let $\mathcal{G}_0$ be a $\{0,1\}$ valued random variable, and for each $e = (w, wi) \in E_*$ with $v \subseteq w$, let $\mathcal{G}_e$ be a $\{0,1\}$ random variable, independent both of each other and of $\mathcal{G}_0$. Suppose that for all $w \in W_*$ and $i \in S$, the distribution of $\mathcal{G}_{(w,wi)}$ depends only on $|w|$.*

The GWVE tree with initial node v, initial value $\mathcal{G}_0$ and edge values $\mathcal{G}_e$ is defined to be the set

$$\mathcal{T} = \left\{ vw \in W_* \,;\, w = w_1 \dots w_n \in W_* \text{ and } \mathcal{G}_0 \prod_{j=1}^{n} \mathcal{G}_{(vw_1 \dots w_{j-1}, vw_1 \dots w_j)} = 1 \right\}. \qquad (6.3.1)$$

If $\mathcal{T}$ is non-empty then we define $\overline{\mathcal{T}}$ to be the set of boundary points of $\mathcal{T}$. If $\mathcal{T}$ is empty then we set $\overline{\mathcal{T}} = \{v\}$.

Note that GWVEs trees are connected subsets of W_*. If $\mathcal{T}$ is a GWVE tree and $w, wi \in \mathcal{T}$, then we say w is the parent of wi, or equivalently that wi is an offspring (or child) of w. Thus the offspring of $w \in \mathcal{T}$ are some subset of $\{wi \in \mathcal{T} \,;\, i \in S\}$.

The GWVE associated to the GWVE tree $\mathcal{T}$ is $n \mapsto |\mathcal{T}_n|$ where $\mathcal{T}_n = \{vw \in \mathcal{T} \,;\, |w| = n\}$, for $n \in \mathbb{N}_0$. The offspring distribution of $w \in \mathcal{T}_n$ is defined to be the distribution of the $\mathbb{N}_0$ valued random variable $\sum_{i=1}^{\mathcal{S}} \mathcal{G}_{w,wi}$. Since the distribution of $\mathcal{G}_{(w,wi)}$ depends only on $|w|$, it is trivial to see that $n \mapsto |\mathcal{T}_n|$ is a GWVE.

Remark 6.3.2 *According to Definition 6.3.1, each edge $e \in E_*$ chooses independently whether $\mathcal{G}_e = 0, 1$. Hence, our GWVE trees can be considered as an inhomogeous per-*

colation on the tree W_*; to describe the corresponding site percolation we say that the site $w \in W_*$ is open if and only if $w \in \mathcal{T}$.

If $\mathcal{T}$ is a GWVE tree with initial node v, then

$$\mathcal{C}_v = \{vw\,;\, w \in W_*,\ \text{there is a path along open nodes connecting } v \text{ to } vw\}$$

is the percolation cluster connected to v. Degeneracy of the GWVE associated to $\mathcal{T}$ is equivalent to $|\mathcal{C}_v| < \infty$. In keeping with our references (e.g. Jirina 1976, Lyons 1992), we tend to use terminology from branching processes rather than percolation.

It follows from Lemma 6.2.1 that if $\mathcal{T}$ is finite and non-empty we have

$$|\overline{\mathcal{T}}| = (\mathcal{S} - 1)|\mathcal{T}| + 1. \tag{6.3.2}$$

Since we defined $\overline{\mathcal{T}}$ to be the initial node of $\mathcal{T}$ when $\mathcal{T}$ is empty, in fact (6.3.2) holds for all finite GWVE trees. Similarly, it follows from Lemma 6.2.1 that if $\mathcal{T}$ and $\mathcal{T}'$ are two finite GWVE trees and $\mathcal{T} \subseteq \mathcal{T}'$ then

$$|\overline{\mathcal{T}}| \leq |\overline{\mathcal{T}'}|. \tag{6.3.3}$$

6.4 GWVE trees and the Segregated Λ-coalescent

The connection to our model is as follows. For each $w \in W_*$, recall that $\mathcal{E}_w$ denotes the clock for the complex K_w which was defined in (5.1.1).

Definition 6.4.1 *For $t > 0$, $v \in W_*$ and $n \in \mathbb{N}_0$ define*

$$\mathscr{B}_n^t(v) = \left\{vw \in W_*\,;\, |w| = n \text{ and for all } j = 0, \ldots, n,\ \mathcal{E}_{vw_1 \ldots w_j} > t\right\}$$
$$\mathscr{B}^t(v) = \bigcup_{n \in \mathbb{N}_0} \mathscr{B}_n^t(v)$$

We write $\mathscr{B}_i^t(\emptyset) = \mathscr{B}_i^t$ and $\mathscr{B}^t(\emptyset) = \mathscr{B}^t$, as well as

$$B_n^t = |\mathscr{B}_n^t|.$$

Note that $\mathscr{B}_n^t(v) = \{w \in \mathscr{B}^t(v)\,;\, |w| = |v| + n\}$, in agreement with the notation ($\mathcal{T}$ and $\mathcal{T}_n$) of Section 6.3. The definitions of $\mathscr{B}_n^t$ and B_n^t given here agree with those of Section 1.2.

Lemma 6.4.2 *For each $t > 0$ and $v \in W_*$, $\mathscr{B}^t(v)$ is a GWVE tree with initial point v, initial value $\mathbb{1}\{\mathcal{E}_v > t\}$ and edge values given for $w \in W_*$ and $i \in S$ by*

$$\mathcal{G}_{(vw,vwi)} = \mathbb{1}\{\mathcal{E}_{vwi} > t\}.$$

The offspring distribution of $vw \in \mathscr{B}_n^t(v)$ is Binomial with $\mathcal{S}$ trials and success probability $e^{-tr_{|v|+n+1}}$.

PROOF: If $vw \in W_*$ and $|w| = n$,

$$\mathbb{1}\{\text{for all } j = 0, \dots n, \mathcal{E}_{vw_1 \dots w_n} > t\} = \mathbb{1}\{\mathcal{E}_v > t\} \prod_{j=1}^{n} \mathbb{1}\{\mathcal{E}_{vw_1 \dots w_n} > t\}. \tag{6.4.1}$$

Since the terms on the right hand side of (6.4.1) are independent and the distribution of $\mathcal{E}_w$ depends only on $|w|$, by comparing the right hand side of (6.4.1) to (6.3.1) we have that $\mathscr{B}^t(v)$ is a GWVE tree with the required parameters.

For $w \in W_*$, note that $e^{-tr_{|vw|}}$ is the probability that $\mathcal{E}_{vw} > t$. Hence, if $vw \in \mathscr{B}_n^t(v)$ then $|w| = n$ and the (conditional) probability that $wi \in \mathscr{B}_{n+1}^t(v)$ is just $e^{-tr_{|v|+n+1}}$. For all $w \in W_*$ and $i, j \in S$, the clocks $\mathcal{E}_{wi}$ and $\mathcal{E}_{wj}$ are independent if $i \neq j$, which gives the required offspring distribution. ■

We can now connect $\mathcal{D}_t$ and A_t to the GWVEs.

Lemma 6.4.3 *For each $t > 0$, almost surely,*

$$\mathcal{D}_t = \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathscr{B}_n^t} K_w.$$

PROOF: Consider $\omega \in \mathcal{C}$, and recall $\mathbb{P}[\mathcal{C}] = 1$ by Lemma 4.2.5. Fix $t > 0$.

Suppose first that $x \in \mathcal{D}_t$. Then for all $w \in W_*$ such that $x \in K_w$, $M_{(0,t] \times \{w\}} = \emptyset$. If $x \in K_w$ and $K_w \subseteq K_{w'}$ then $x \in K_{w'}$ and $M_{(0,t] \times \{w'\}} = \emptyset$. Hence, if $x \in K_w$ then $w \in \mathscr{B}_{|w|}^t$.

By [$\mathscr{K}$1], for all $n \in \mathbb{N}$ there is precisely one $w \in W_n$ such that $x \in K_w$. Thus $x \in \bigcup_{w \in \mathscr{B}_n^t} K_w$ for all n, and hence $\mathcal{D}_t \subseteq \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathscr{B}_n^t} K_w$.

Now suppose that $x \in \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathscr{B}_n^t} K_w$. The argument above essentially works in reverse. Then $x \in \bigcup_{w \in \mathscr{B}_n^t} K_w$ for all n, from which it follows that if $x \in K_w$ we must have $M_{(0,t] \times \{w\}} = \emptyset$. Thus $(E_m^{x,0,t}) = \emptyset$ and $x \in \mathcal{D}_t$. ■

Lemma 6.4.4 *For all $t > 0$, almost surely,*

$$A_t = \left\{ K_w \,;\, w \in \overline{\mathscr{B}^t} \right\}.$$

PROOF: If $w \in \overline{\mathscr{B}^t}$ and $w = w_1 \dots w_n$ for $w_j \in S$, then $\mathcal{E}_{w_1 \dots w_j} > t$ for all $j < n$ and $\mathcal{E}_w \leq t$. Hence, for all $x \in K_w$, $(E^{x,0,t})$ is a finite sequence, with its final element given by

$$(u, w, p) \in M \text{ such that } u = \max\{s \,;\, (s, w, p) \in M\}.$$

Hence $X_{0,t}(x) = p$ for all $x \in K_w$. If $y \notin K_w$ then, since $\mathcal{E}_{w_1 \ldots w_j} > t$ for all $j < n$, $X_{0,t}(y) \notin K_w$. Hence $K_w \in A_t$.

Conversely, if w is such that $K_w \in A_t$, then $X_{0,t}(K_w)$ is a single point, so $\mathcal{E}_{w_1 \ldots w_j} \leq t$ for at least one $j \leq n$. If $K_w = K$ then we are done, since this means $\mathcal{E}_\emptyset \leq t$. If $K_w \neq K$ then for all $w' \in W_*$ such that $K_w \subseteq K_{w'}$ and $K_w \neq K_{w'}$, there is $y \in K_{w'} \setminus K_w$. Since $y \notin K_w$ we must have $\mathcal{E}_w > t$. In particular, $\mathcal{E}_{w_1 \ldots w_j} > t$ for all $j < n$. Hence $\mathcal{E}_w \leq t$, so as $w \in \overline{\mathscr{B}^t}$.

■

Lemmas 6.4.3 and 6.4.4 will be crucial to our analysis of the Segregated Λ-coalescent in all of the following chapters.

It follows from Lemmas 6.4.2 and 6.1.1 that

$$\mathbb{E}\left[B_n^t(v)\right] = \mathcal{S}^n \exp\left(-t \sum_{j=0}^{n} r_{|v|+j}\right) \tag{6.4.2}$$

$$\operatorname{var}\left[B_n^t(v)\right] = \mathbb{E}\left[B_n^t(v)\right]^2 \sum_{l=0}^{n} \frac{1 - e^{-t r_{|v|+l}}}{\mathcal{S}^l \exp\left(-t \sum_{j=0}^{l} r_{|v|+j}\right)}. \tag{6.4.3}$$

6.5 Two results using GWVE trees

In this section we establish two useful results about the Segregated Λ-coalescent. These results are important in their own right and we will use them in Chapters 8-10 (although not in Chapter 7).

First, we give a formula for the expectation of $|A_t|$. If $\mathcal{D}_t = \emptyset$ then there are no singletons and the only blocks in the Segregated Λ-coalescent are precisely the elements of A_t. Thus, the following Lemma is an important tool for determining the rate at which the Segregated Λ-coalescent comes down from infinity. Unfortunately, in all but the simplest cases (see Lemma 8.2.1) it does not seem possible to obtain a closed form of the formula.

Lemma 6.5.1 *Suppose that A_t is almost surely finite. Then,*

$$\mathbb{E}[|A_t|] = 1 + (\mathcal{S} - 1) \sum_{n=0}^{\infty} \mathcal{S}^n e^{-t \sum_0^n r_j}. \tag{6.5.1}$$

PROOF: By definition of A_t we have $|A_t| = |\overline{\mathscr{B}^t}|$, which is finite by our hypothesis. By (6.3.2) (which applies since $|\overline{\mathscr{B}^t}|$ is almost surely finite) we have

$$|\overline{\mathscr{B}^t}| = (\mathcal{S} - 1)|\mathscr{B}^t| + 1.$$

Noting $|\mathscr{B}^t| = \sum_0^\infty B_n^t$ and using (6.4.2) we thus have

$$\begin{aligned}\mathbb{E}\left[|\overline{\mathscr{B}^t}|\right] &= (\mathcal{S}-1)\sum_{n=0}^{\infty}\mathbb{E}\left[B_n^t\right]+1\\ &= 1+(\mathcal{S}-1)\sum_{n=0}^{\infty}\mathcal{S}^n e^{-t\sum_0^n r_j}.\end{aligned}$$

This completes the proof. ■

Secondly, we give a monotonicity result for X in terms of its coagulation rates. Recall that the number of blocks in the Λ-coalescent decreases as the rate of coalescence increases. To be precise, if Λ and $\widetilde{\Lambda}$ are two measures on $[0,1]$ and for all measurable $A \subseteq [0,1]$ we have $\Lambda(A) \leq \widetilde{\Lambda}(A)$, then the associated Λ-coalescents Π and Π_t can be coupled in such a way that

$$|\widetilde{\Pi_t}| \leq |\Pi_t|$$

for all $t \geq 0$. We now give an equivalent result for the Segregated Λ-coalescent.

Let $(\widetilde{r_n})_{n\in\mathbb{N}_0}$ be a sequence in $[0,\infty)$ such that

$$\widetilde{r_n} \geq r_n \tag{6.5.2}$$

for all $n \in \mathbb{N}_0$.

Let $\widetilde{X}$ be the Segregated Λ-coalescent associated to the coagulation rates $\widetilde{r}_n$ and the measure $\widetilde{\mathcal{U}} = \mathcal{U}$. More precisely, if $\widetilde{M}$ is the Poisson point process from which $\widetilde{X}$ is defined, by (6.5.2) and the standard theory of Poisson point processes, we can couple M and $\widetilde{M}$ in such a way that $M \subseteq \widetilde{M}$. Thus X and $\widetilde{X}$ are defined on the same probability space which (with minor abuse of notation) we will denote by $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $\widetilde{A_t}$ and $\widetilde{D_t}$ be the dust and non-singleton blocks of $\widetilde{X}$ at time t. Let $\widetilde{\mathscr{B}^t}$ denote the GWVE tree associated to X at time t, and let $\widetilde{\mathcal{E}_w}$ be the exponential clock for w which is associated to $\widetilde{M}$.

Lemma 6.5.2 *Almost surely, for all $t \geq 0$, $|\widetilde{A_t}| \leq |A_t|$ and $\widetilde{\mathcal{D}_t} \subseteq \mathcal{D}_t$.*

Proof: Since $M \subseteq \widetilde{M}$, for all $w \in W_*$ we have $\mathcal{E}_w \leq \widetilde{\mathcal{E}_w}$ and thus also $\widetilde{\mathscr{B}^t} \subseteq \mathscr{B}^t$. By Lemma 6.4.3 we then have $\widetilde{\mathcal{D}_t} \subseteq \mathcal{D}_t$. Similarly, by (6.3.3) and Lemma 6.4.4 we have $|\widetilde{A_t}| \leq |A_t|$. ■

Chapter 7

Proof of the phase classification

In this Chapter we prove Theorem 5.2.1.

7.1 Some preliminary results

In this section we give some results, working in the direction of Theorem 5.2.1, using Fubini's theorem and the spatial structure of K.

Lemma 7.1.1 *For all $t > 0$,*

- *If $\sum r_n = \infty$ then $\mathbb{P}[\lambda_t(\mathcal{D}_t) = 0] = 1$.*
- *If $\sum r_n < \infty$ then $\mathbb{P}[\lambda_t(\mathcal{D}_t) > 0] > 0$.*

PROOF: Fix $t > 0$. By Fubini's theorem, which applies by Lemma 4.5.3,

$$\mathbb{E}[\lambda(\mathcal{D}_t)] = \mathbb{E}\left[\int_K \mathbb{1}\{x \in \mathcal{D}_t\}\lambda(dx)\}\right] = \int_K \mathbb{P}[x \in \mathcal{D}_t]\,\lambda(dx). \tag{7.1.1}$$

By Lemma 4.2.3, if $\sum r_n < \infty$ then for each $x \in K$ there is positive probability that x was not involved in any reproduction events during $[0, t]$. Hence $\mathbb{P}[x \in \mathcal{D}_t] > 0$, so that $\mathbb{P}[\lambda(\mathcal{D}_t) > 0] > 0$ by (7.1.1). Also by Lemma 4.2.3, if $\sum r_n = \infty$ then each $x \in K$ has almost surely been involved in some reproduction event during $[0, t]$ and thus $\mathbb{P}[x \in \mathcal{D}_t] = 0$. By (7.1.1), $\mathbb{P}[\lambda(\mathcal{D}_t) = 0] = 1$. ■

Lemma 7.1.2 *Suppose A_t is finite. Then, almost surely, either $\mathcal{D}_t = \emptyset$ or, for some $w \in W_*$, $K_w \subseteq \mathcal{D}_t$. In the latter case, $\lambda_t(\mathcal{D}_t) > 0$.*

PROOF: Suppose that A_t is finite and $\mathcal{D}_t \neq \emptyset$. By Lemma 5.5.1 $\mathscr{A}_t$ is finite, say $\mathscr{A}_t = \{y(i)\,;\, i = 1, \ldots, N\} \subseteq K$. Also by Lemma 5.5.1, for each $y(i) \in \mathscr{A}_t$ there is some $x(i) \in K$

such that $X_{0,t}^{-1}(y(i)) = K_{w_1}^{x(i),0,t}$. Let $n = \max\{|w_1^{x(i),0,t}|\,;\, i = 1,\ldots,N\}$. By [$\mathscr{K}$1], each $K_{w_1}^{x(i),0,t}$ can be written as the union of finitely many K_w such that $w \in W_n$. Since $\mathcal{D}_t \neq 0$ we must have $\bigcup_i K_{w_1}^{x(i),0,t} \subset K$, and thus there is some $w \in W_n$ such that $K_w \subseteq \mathcal{D}_t$. By Lemma 4.1.2 we then have $\lambda_t(\mathcal{D}_t) > 0$. ■

Lemma 7.1.3 *Suppose* $\sum \mathcal{S}^n r_n = \infty$. *Then* $\mathbb{P}[\exists w \in W_*, K_w \subseteq \mathcal{D}_t] = 0$.

PROOF: Fix $t > 0$. Since W_* is countable, the lemma follows if we can show that $\mathbb{P}[K_w \subseteq \mathcal{D}_t] = 0$ for an arbitrary $w \in W_*$. So fix $w \in W_*$, and set $n = |w|$. The rate at which K_w is affected by reproduction events is

$$\begin{aligned}\int_{W^*\times K} \mathbb{1}\{K_w \cap K_{w'} \neq \emptyset\}\mathcal{P}(dw', dp) &= \sum_{w'\in W_*} r_{|w|}\mathbb{1}\{K_w \cap K_{w'} \neq \emptyset\} \\ &\geq \sum_{w'\in W_*} r_{|w|}\mathbb{1}\{K_{w'} \subseteq K_w\}\end{aligned}$$

Now, by [$\mathscr{K}$1], $K_{w'} \subseteq K_w$ if and only if $w' = wv$ for some $v \in W_*$. Hence,

$$\sum_{w'\in W_*} r_{|w|}\mathbb{1}\{K_{w'} \subseteq K_w\} = \sum_{v\in W_*} r_{|wv|} = \frac{1}{\mathcal{S}^n}\sum_{m=n}^{\infty} \mathcal{S}^m r_m = \infty$$

It follows immediately that (with probability one) K_w is affected by a reproduction event during $(0, t]$ for any $t > 0$. Hence $\mathbb{P}[K_w \subseteq \mathcal{D}_t] = 0$. ■

7.2 Conditioning on non-degeneracy

In this section we establish a result which we will need in the proof of Lemma 7.3.1, to deal with cases where K is not completely segregated.

Let $I(\mathscr{B}^t)$ be the connected subset of W_* which is made up of the infinite paths of $\mathscr{B}^t$, that is

$$I(\mathscr{B}^t) = \{w \in W_*\,;\, w \text{ is a node on some infinite path of } \mathscr{B}^t\}. \tag{7.2.1}$$

For each $w \in W_*$, let

$$C_w = \{wi\,;\, i \in S, wi \in I(\mathscr{B}^t)\}$$

be the set of children of w in $I(\mathscr{B}^t)$.

For the duration of this section we fix $t > 0$. Let $\mathbb{P}_n$ be the probability measure defined by

$$\mathbb{P}_n[A] = \frac{\mathbb{P}\left[A \cap \{B_n^t = 1\}\right]}{\mathbb{P}\left[B_n^t = 1\right]}$$

and note that this is well defined by Lemma 6.4.2. Let

$$q_n = \mathbb{P}_n\left[\exists m \in \mathbb{N}, B_m^t = 0\right].$$

If $q_0 < 1$ then B_n^t is non-degenerate with positive probability and we can condition on the event

$$Q = \{\forall n, B_n^t \geq 1\} = \{I(\mathscr{B}^t) \neq \emptyset\}. \tag{7.2.2}$$

Let $\mathbb{P}_Q$ be the probability measure defined by

$$\mathbb{P}_Q[A] = \frac{\mathbb{P}[A \cap Q]}{\mathbb{P}[Q]}.$$

Remark 7.2.1 *We need language to distinguish between the distributions of random variables under $\mathbb{P}$ and under $\mathbb{P}_Q$. We do so by referring to (the distribution of) 'Z under Q' or 'Z conditional on Q'. We will need to make this distinction only during this section and in (the proof of) Lemma 7.3.1.*

The purpose of this section is to describe the set of infinite paths of $\mathscr{B}^t$ under $\mathbb{P}_Q$. Theorem 4.10 of Lyons (1992) describes the 'genealogy of $(B_n^t)_{n\in\mathbb{N}}$ conditional on Q', by which Lyons means the topological structure of the connected graph $\mathscr{B}^t$ under the measure $\mathbb{P}_Q$. To be precise Lyons does not embed the genealogical trees of GWVEs into W_* (or any other space) and instead considers them as connected graphs in their own right.

Remark 7.2.2 *The classical results which describe Galton-Watson processes conditioned to be degenerate/non-degenerate are well known, see pages 49 and 52 of Athreya and Ney (1972).*

Theorem 4.10 of Lyons (1992) tells us the distribution of $|C_w|$ under $\mathbb{P}_Q$, which we now describe. Note first that $\mathbb{P}_Q\left[\emptyset \in I(\mathscr{B}^t)\right] = 1$ by definition of Q. Under $\mathbb{P}_Q$, $\{|C_w|\,;\, w \in W_*\}$ are independent random variables and if $|w| = n \in \mathbb{N}_0$ then $|C_w|$ has generating function

$$f_n^*(s) = \frac{1}{1-q_{n-1}}\left[f_n\big(q_n + (1-q_n)s\big) - q_{n-1}\right] \tag{7.2.3}$$

where f_n is the generating function of $|C_w|$ under $\mathbb{P}$, which by Lemma 6.4.2 is $f_n(s) = (1 - e^{tr_n} + se^{tr_n})^{\mathcal{S}}$. In order to understand $I(\mathscr{B}^t)$ under $\mathbb{P}_Q$, we will need to discover precisely which set C_w of children for each w has been chosen by the conditioning.

Remark 7.2.3 *A short calculation, using that*

$$q_n = \lim_{N\to\infty} f_{n+1} \circ \ldots \circ f_{n+N}(0), \tag{7.2.4}$$

shows that $\mathbb{P}_Q\left[|C_w| = 0\right] = f_{|w|}(0) = 0$*, which means that* $|C_w| \geq 1$ *for all* $w \in I(\mathscr{B}^t)$*, under* $\mathbb{P}_Q$*.*

Recall that $\mathcal{E}_w$ and $\mathcal{E}_{w'}$ are independent for $w \neq w'$, and identical in distribution if $|w| = |w'|$. Hence, conditional on $w \in I(\mathscr{B}^t)$ having precisely $k \neq 0$ offspring in $I(\mathscr{B}^t)$, the distribution of C_w is uniform on the set

$$\Big\{A \subseteq \{wi\,;\, i \in S\}\,;\, |A| = k\Big\}.$$

In words, the above is the set of subsets of $\{wi\,;\, i \in S\}$ which have precisely k elements.

We can now describe $\mathscr{B}^t$ in full, which we record in the following lemma.

Lemma 7.2.4 *Suppose that* $q_0 < 1$*. The distribution of* $I(\mathscr{B}^t)$ *under* $\mathbb{P}_Q$ *is equal to the distribution of the tree* $\mathcal{T}$ *which is described as follows.*

- *First, for each* $w \in W_n$ *sample independent random variables* k_w *with generating function given by* (7.2.3)*. Then for each* $w \in W_*$ *for which* $k_w \neq 0$*, sample independent (both of each other and of* $\{k_u\,;\, u \neq w\}$*) random variables* A_w*, uniformly on*

$$\Big\{A \subseteq \{wi\,;\, i \in S\}\,;\, |A| = k_w\Big\}.$$

- *Include* $\emptyset \in \mathcal{T}$*. Then, working inductively outwards from* $\emptyset$ *along the tree* W_**: If* $w \in \mathcal{T}$ *then include also* $A_w \subseteq \mathcal{T}$*.*

It follows immediately from this description that $I(\mathscr{B}^t)$ is a GWVE tree under $\mathbb{P}_Q$, with offspring distributions given by (7.2.3).

7.3 Dust and GWVEs

In this section we build on the results of Lemmas 6.4.3 and 6.4.4.

Lemma 7.3.1 *Let* $t > 0$*. Then, almost surely,* $\mathcal{D}_t = \emptyset$ *if and only if* $\exists n \in \mathbb{N}, B_n^t = 0$*.*

PROOF: Fix $t > 0$. Suppose first that for some (random) $n \in \mathbb{N}$, $B_n^t = 0$. Then, by Lemma 6.4.3, $\mathcal{D}_t = \emptyset$.

We approach the converse argument as follows. Recall that $q_0 = \mathbb{P}\left[\exists n \in \mathbb{N}, B_n^t = 0 \,|\, B_0^t = 1\right]$. If $q_0 = 1$ then we have nothing to prove, so we may assume $q_0 < 1$.

Remark 7.3.2 *On Q, it is easy to see that there exists a sequence $(w(n))_{n\in\mathbb{N}}$ such that $w(n) \in \mathscr{B}_n^t$ and $K_{w(n)} \supseteq K_{w(n+1)}$. If K is completely segregated then, by Lemma 4.1.6 we have that each K_w is closed, from which it follows by completeness of K that $\cap_{n\in\mathbb{N}} K_{w(n)}$ is non-empty. Since $\cap_{n\in\mathbb{N}} K_{w(n)} \subseteq \mathcal{D}_t$ our proof would be complete. This argument has appeared several times in the random fractals literature e.g. Lemma 8 of Durand (2009).*

However, if K is not completely segregated then K_w might not be closed. In our more general set up we must invoke [$\mathscr{K}$5] and use Lemma 7.2.4 to prove (a close equivalent of) $\cap_{n\in\mathbb{N}} K_{w(n)} \neq \emptyset$.

For the remainder of the proof we condition on Q and work under the measure $\mathbb{P}_Q$, without further comment. Recall that the corresponding distribution of $I(\mathscr{B}^t)$ was described in Lemma 7.2.4.

The set $I(\mathscr{B}^t)$ consists entirely of (nodes of) infinite paths of W_*. By definition, each $v \in I(\mathscr{B}^t)$ has at least one infinite path of $I(\mathscr{B}^t)$ beginning at v. In particular, each $v \in I(\mathscr{B}^t)$ has at least one child in $I(\mathscr{B}^t)$. With this in mind, we define a *random* infinite path $\mathscr{U}$ of $I(\mathscr{B}^t)$ as follows. Pick the first edge of $\mathscr{U}$ uniformly from the set

$$\{(\emptyset, i)\,;\, i \in I(\mathscr{B}^t)\}.$$

Working inductively, if $(w, wi) \in \mathscr{U}$ then we pick the next edge of $\mathscr{U}$ uniformly from the set

$$\{(wi, wij)\,;\, j \in S, wij \in I(\mathscr{B}^t)\}.$$

Let the set $(w(n))_{n\in\mathbb{N}_0}$ be the (ordered) set of nodes on the infinite path $\mathscr{U}$. Note that $|w(n)| = n$.

By [$\mathscr{K}$5], for each $w \in W_*$ the set

$$\mathcal{J}_w = \{i \in S\,;\, \overline{K_{wi}} \subseteq K_w\}$$

has at least one element. By definition of $\mathscr{U}$ and Lemma 7.2.4, the random variables $(\mathbb{1}\{i \in \mathcal{J}_{w(n)}\})_{n\in\mathbb{N}_0, 1\leq i\leq \mathcal{S}}$ are independent. Since $\mathcal{J}_w$ has at least one element, also by definition of $\mathscr{U}$ and Lemma 7.2.4,

$$\mathbb{P}_Q\left[w(n_i) \in \mathcal{J}_{w(n)}\right] \geq \frac{1}{\mathcal{S}}.$$

By the Borel-Cantelli lemma, with probability one there is an infinite subsequence $(n_i)_{i\in\mathbb{N}_0}$ of $\mathbb{N}_0$ such that $\lim_{i\to\infty} n_i = \infty$ and $\overline{K_{w(n_i+1)}} \subseteq K_{w(n_i)}$ for all i.

Since $w(n+1)$ is a child of $w(n)$, for all n, the sequence $K_{w(n)}$ is a decreasing sequence of sets. Let x_n be some point in $K_{w(n)}$, then by completeness of K and [$\mathscr{K}$2] we have that there exists $x \in K$ such that $x_n \to x$. Note that $x \in \cap_{i\in\mathbb{N}_0}\overline{K_{w(n_i)}}$.

For all $n \in \mathbb{N}$ there is some $i \in \mathbb{N}$ such that $n_i < n$, so that $K_{w(n)} \subseteq \overline{K_{w(n_i)}}$. Hence, by Lemma 6.4.3,

$$x \in \bigcap_{i=0}^{\infty} \overline{K_{w(n_i)}} \subseteq \bigcap_{n=0}^{\infty} K_{w(n)} \subseteq \mathcal{D}_t.$$

We thus have $\mathbb{P}_Q[\mathcal{D}_t \neq \emptyset] = 1$, which completes the proof. ■

Lemma 7.3.3 *Let $t > 0$. Almost surely, if $\mathcal{D}_t = \emptyset$ then A_t is finite.*

PROOF: If $\mathcal{D}_t$ is empty then, by Lemma 7.3.1 there is almost surely some $n \in \mathbb{N}$ such that $B_n^t = \emptyset$. Hence also $B_m^t = 0$ for all $m \neq n$. By Lemma 6.4.4, A_t is finite. ■

Lemma 7.3.4 *Let $t > 0$. Then $\mathbb{P}\big[\lim_n B_n^t = \infty$ or $\exists n, B_n^t = 0\big] = 1$. Further, almost surely, $\mathcal{D}_t \neq \emptyset$ if and only if $\lim_{n\to\infty} B_n^t = \infty$.*

PROOF: To prove the first statement, we use the result of Theorem 1 in Jagers (1974), which is a restatement (with minor correction) of a result in Church (1967).

The probability of a individual at stage n in the process B_t having exactly one offspring is given by $p_{n1}^t = \mathcal{S}e^{-r_n t}(1-e^{-r_n t})^{\mathcal{S}-1}$. Note that for $a \in [0,1]$ and $n \geq 1$, $a(1-a)^n \leq \frac{1}{n+1}(1-\frac{1}{n+1})^n$. Since $\mathcal{S} \geq 2$ we have $p_{n1}^t \leq (1-1/\mathcal{S})^{\mathcal{S}-1} < 1$. Hence $\sum_n (1-p_{n1}^t) = \infty$, and from Jagers (1974) we have

$$\mathbb{P}\left[\lim_{n\to\infty} B_n^t = \infty \text{ or } \exists n \geq 0, B_n^t = 0\right] = 1.$$

It follows immediately from this and Lemma 7.3.1 that $\lim_{n\to\infty} B_n^t = \infty$ is almost surely equivalent to $\mathcal{D}_t \neq \emptyset$. ■

Lemma 7.3.5 *Let $t > 0$. If $\sum r_n < \infty$ then $\mathbb{P}[\mathcal{D}_t = \emptyset$ or $\lambda(\mathcal{D}_t) > 0] = 1$.*

PROOF: The process $n \mapsto B_n^t / \mathbb{E}\left[B_n^t\right]$ is a discrete parameter, non-negative martingale. By the martingale convergence theorem there is some random variable L^t such that $\frac{B_n^t}{\mathbb{E}[B_n^t]} \to L^t$ almost surely.

Recall that in (6.4.2) we gave a formula for $\mathbb{E}\left[B_n^t\right]$. Since $\sum r_n < \infty$,

$$\mathbb{E}\left[B_n^t\right] \geq \mathcal{S}^n \exp\left(-t\sum_0^\infty r_j\right) = C\mathcal{S}^n. \tag{7.3.1}$$

where $C = C(t) > 0$. In the language of Biggins and D'Souza (1992), (7.3.1) means that B^t is uniformly supercritical. Since the offspring distribution of B^t is uniformly bounded (by $\mathcal{S}$), Theorem 2 of Biggins and D'Souza (1992) applies. In our notation this means that

$$\left\{B_n^t \to \infty\right\} = \left\{L^t > 0\right\}. \tag{7.3.2}$$

Now, suppose $\omega \in \mathcal{C}$ and that $\mathcal{D}_t \neq \emptyset$. By Lemma 6.4.3, for all $n \in \mathbb{N}$ we have $B_n^t = |\mathscr{B}_n^t| \geq 1$. By the first part of Lemma 7.3.4 it follows that (almost surely) $B_n^t \to \infty$ as $n \to \infty$. By (7.3.2), $\lim_{n\to\infty} \frac{B_n^t}{\mathbb{E}[B_n^t]} > 0$. From this and (7.3.1),

$$\liminf_{n\to\infty} \frac{B_n^t}{C\mathcal{S}^n} > 0 \tag{7.3.3}$$

where the lim inf could potentially be infinite. In fact, though, $B_n^t \leq B_0^t\mathcal{S}^n \leq \mathcal{S}^n$ so the lim inf in (7.3.3) is finite. We write

$$l = \liminf_{n\to\infty} \frac{B_n^t}{C\mathcal{S}^n}$$

where $l \in (0,\infty)$ (note l is random). Then there exists $N \in \mathbb{N}$ such that for all $n > N$, $\frac{B_n^t}{C\mathcal{S}^n} \geq l/2$. So for all $n > N$ we have $B_n^t \geq \frac{Cl}{2}\mathcal{S}^n$.

Note that the sets $\bigcup_{w\in\mathscr{B}_n^t} K_w$ are decreasing as n increases. By Lemma 6.4.3, $\lambda(\mathcal{D}_t) = \lim_n \lambda\left(\bigcup_{w\in\mathscr{B}_n^t} K_w\right)$. Recall that $\lambda(K_w) > 0$ and $\lambda(K_w)$ depends only on $|w|$ (by Lemma 4.1.2 and [$\mathscr{K}$3] respectively). Hence $\lambda(K_w) = \frac{\lambda(K)}{\mathcal{S}^n}$. By [$\mathscr{K}$1],

$$\lambda\left(\bigcup_{w\in\mathscr{B}_n^t} K_w\right) = \sum_{w\in\mathscr{B}_n^t} \lambda(K_w) = |\mathscr{B}_n^t|\lambda(K_w) = \frac{B_n^t}{\mathcal{S}^n}\lambda(K).$$

Thus, for $n > N$, $\lambda\left(\bigcup_{w\in\mathscr{B}_n^t} K_w\right) \geq Cl\lambda(K)$. Hence, $\lambda(\mathcal{D}_t) > 0$. ■

Remark 7.3.6 *Whilst results like* (7.3.2) *are well understood for the classical supercritical Galton-Watson processes, it turns out that they are not true in general for supercritical GWVEs. For some GWVEs,* (G_n)*, it holds that* $\mathbb{E}[G_n] \to \infty$ *and* $\left\{\lim \frac{G_n}{\mathbb{E}[G_n]} > 0\right\}$ *is a proper subset of* $\{G_n \to \infty\}$*. The reason for this is that (as we mentioned in Section 5.4) GWVEs can have multiple rates of asymptotic growth. The result we used above from Biggins and D'Souza (1992) shows that, when* $\sum r_n < \infty$*,* B^t *has only one rate of growth.*

7.4 Degeneracy of the GWVEs

To apply Lemmas 7.3.1 and 7.3.4 we need a characterisation of the extinction/explosion of B^t. As we mentioned in Section 5.4, we use degeneracy conditions from Jirina (1976). Let us write $m_n^t = \mathbb{E}\left[B_n^t\right]$. From (6.4.2) we have

$$m_n^t = \exp\left(n\log\mathcal{S} - t\sum_{j=0}^{n} r_j\right). \tag{7.4.1}$$

Clearly $m_n^t \in (0,\infty)$. The quantity $\inf_n m_n^t$ will feature in these conditions, along with

$$g^t = \sum_{n=1}^{\infty} \frac{(1-e^{-r_{n+1}t})^{\mathcal{S}} + \mathcal{S}e^{-r_{n+1}t} - 1}{\mathcal{S}e^{-r_{n+1}t}\, m_n^t}. \tag{7.4.2}$$

For all $x > 0$ and $n \in \mathbb{N}$, $n \geq 2$, it holds that $(1-x)^n + nx - 1 > 0$, and hence $g^t \in (0,\infty]$.

Lemma 7.4.1 (Jirina 1976) $\mathbb{P}\left[\exists n \in \mathbb{N}, B_n^t = 0\right] < 1$ *if and only if both* $\inf m_n^t > 0$ *and* $g^t < \infty$.

Proof: Lemma of Jirina (1976) 1.1 tells us that if $\inf_n m_n^t = 0$, then $\mathbb{P}\left[\exists n \in \mathbb{N}, B_n^t = 0\right] = 1$. Our offspring distributions are binomial with $\mathcal{S}$ trials, which is sufficient for Theorem 2.3 of Jirina (1976) to apply. From this we obtain that, if $\inf_n m_n^t > 0$, then $\mathbb{P}\left[\exists n \in \mathbb{N}, B_n^t = 0\right] = 1$ if and only if $g^t = \infty$. ■

The remainder of Section 7.4 will consist of real analysis, with the objective of pinning down the behaviour of $\inf_n m_n^t$ and g^t.

Lemma 7.4.2 *Suppose that* $v \in (0,\infty)$ *is such that* $\inf_n m_n^v > 0$. *Then for all* $u \in (0,v)$, $\inf_n m_n^u > 0$ *and* $g^u < \infty$.

Proof: Let $\inf_n m_n^v > 0$. Suppose that $\epsilon > 0$ and for infinitely many n we have $\frac{1}{n}\sum_0^n r_j > \frac{\epsilon + \log\mathcal{S}}{v}$. For such n,

$$\begin{aligned} m_n^v &= \left(\exp\left(\log\mathcal{S} - v\frac{1}{n}\sum_0^n r_j\right)\right)^n \\ &\leq \left(\exp\left(\log\mathcal{S} - v\frac{\epsilon+\log\mathcal{S}}{v}\right)\right)^n = e^{-\epsilon n}. \end{aligned}$$

This may not occur for infinitely many n since $\inf m_n^v > 0$. So, we may assume that both $\inf_n m_n^v > 0$, and

$$\limsup_n \frac{1}{n}\sum_0^n r_j \leq \frac{\log\mathcal{S}}{v}. \tag{7.4.3}$$

Let $u \in (0, v)$. Then, for some $\epsilon > 0$ we have $0 < u \limsup_n \frac{1}{n} \sum_0^n r_j < \log \mathcal{S} - \epsilon$. Hence, there exists $N \in \mathbb{N}$ (dependent on ϵ) such that for all $n > N$, $0 < u\frac{1}{n} \sum_0^n r_j < \log \mathcal{S} - \epsilon$. Thus,

$$m_n^u = \left(\exp\left(\log \mathcal{S} - u \frac{1}{n} \sum_0^n r_j \right) \right)^n \geq \left(\exp\left(\epsilon \right) \right)^n,$$

so clearly $\inf m_n^u > 0$. Also

$$g^u \leq \sum_1^\infty \frac{\mathcal{S}}{\mathcal{S} e^{-r_{n+1} u} m_n^u} = \mathcal{S} \sum_1^\infty \frac{1}{m_{n+1}^u} \leq \mathcal{S} \sum_1^\infty \frac{1}{(e^\epsilon)^{n+1}} < \infty$$

as required. ■

Lemma 7.4.3 *The following hold:*

1. $\inf m_n^t = 0$ *for all* $t > 0$ *if and only if* $\limsup_{n\to\infty} \frac{1}{n} \sum_0^n r_j = \infty$.
2. $\inf m_n^t > 0$ *for all* $t > 0$ *if and only if* $\limsup_{n\to\infty} \frac{1}{n} \sum_0^n r_j = 0$.
3. $\exists t_0 \in (0, \infty)$ *such that* $t < t_0 \Rightarrow \inf_n m_n^t > 0$ *and* $t > t_0 \Rightarrow \inf m_n^t = 0$, *if and only if,* $\limsup_{n\to\infty} \frac{1}{n} \sum_0^n r_j \in (0, \infty)$.

   For all (r_n), *precisely one of the above cases occurs.*

PROOF: Since $\mathcal{L} = \limsup_{n\to\infty} \frac{1}{n} \sum_0^n r_j$ exists in $[0, \infty]$ precisely one of $\mathcal{L} = 0$, $\mathcal{L} = \infty$ and $\mathcal{L} \in (0, \infty)$ occurs. We give each case in turn.

CASE 1: Suppose that $\limsup_n \frac{1}{n} \sum_0^n r_j = \infty$. For any $t > 0$, we can pick a subsequence (r_{i_n}) of (r_n) such that for all n, $\frac{1}{i_n} \sum_0^{i_n} r_j > \frac{\log \mathcal{S} + 1}{t}$. Hence $m_{i_n}^t \leq (\exp(-1))^{i_n}$ for all n, and since $i_n \to \infty$ it follows that $\inf_n m_n^t = 0$.

Conversely, if $\limsup_n \frac{1}{n} \sum_0^n r_j = C < \infty$ then for $t = \frac{\log \mathcal{S}}{2C} > 0$ we note that

$$m_n^t = \left(\exp\left(\log \mathcal{S} - \frac{\log \mathcal{S}}{2} \frac{1}{C} \frac{1}{n} \sum_0^n r_j \right) \right)^n.$$

For sufficiently large n, $\frac{1}{n} \sum_0^n r_j \leq \frac{3}{2} C$, and hence for sufficiently large n, $m_n^t \geq \left(\exp\left(\frac{1}{3} \log \mathcal{S} \right) \right)^n$. Hence $\inf_n m_n^t > 0$.

CASE 2: Suppose that $\limsup_n \frac{1}{n} \sum_0^n r_j = 0$ and let $t > 0$. Then, for all sufficiently large n we have $\frac{1}{n} \sum_0^n r_j \leq \frac{1}{2t}$. Hence, for all sufficiently large n, $m_n^t \geq \left(\exp\left(\frac{1}{2} \log \mathcal{S} \right) \right)^n$. Thus $\inf_m m_n^t > 0$.

Conversely, suppose that $\inf_n m_n^t > 0$ for all t. Fixing t, and using the first step of the proof of Lemma 7.4.2, we obtain from (7.4.3) that $\limsup_n \frac{1}{n}\sum_0^n r_j \leq \frac{\log \mathcal{S}}{t}$. However, we have $\inf_n m_n^t > 0$ for all $t > 0$, so $\limsup_n \frac{1}{n}\sum_0^n r_j = 0$.

CASE 3: It follows immediately from the definition of m_n^t that for $t < s$, $m_n^t \geq m_n^s$. Thus, $\inf_n m_n^t \geq \inf m_n^s$. Hence, for each $\mathcal{S}$ and (r_n) there is a unique $t^* \in [0, \infty]$ such that $\inf_n m_n^t > 0$ for $t < t^*$ and $\inf m_n^t = 0$ for $t > t^*$. This case now follows from cases 1 and 2. ■

The above proof makes it clear that we should try to obtain an explicit formula for t^*, which in the critical case is equal to the critical time t_0. We do so in Corollary 8.1.1. The following result is needed to understand the behaviour of $\mathcal{D}_{t_0}$ and A_{t_0} in Lemma 8.1.2.

Lemma 7.4.4 *There exist $C_1, C_2 \in (0, \infty)$ (dependent only upon $\mathcal{S}$) such that for all $t > 0$,*

$$C_1 \sum_n \frac{e^{-r_{n+1}t}}{m_n^t} \leq g^t \leq C_2 \sum_n \frac{e^{-r_{n+1}t}}{m_n^t}.$$

PROOF: Let $f_n : (0, \infty) \to (0, \infty)$ be defined by

$$f_n(x) = (1 - x)^n + nx - 1.$$

It is elementary to show that there exists $C_1, C_2 \in (0, \infty)$ (dependent only upon $\mathcal{S}$) such that for all $x \in (0, 1]$,

$$C_1 x^2 \leq f_{\mathcal{S}}(x) \leq C_2 x^2.$$

Since

$$g^t = \sum_n \frac{f_{\mathcal{S}}(e^{-r_{n+1}t})}{\mathcal{S}e^{-r_{n+1}t}\, m_n^t},$$

the stated result follows. ■

We have now built up enough ammunition to prove Theorem 5.2.1.

7.5 Proof of Theorem 5.2.1

It is elementary to show that Theorem 5.2.1 assigns each possible choice of $\mathcal{S}$ and (r_n) to precisely one phase. Hence, it suffices to show that the criteria given by Theorem 5.2.1 for each phase are sufficient.

Recall that for at least one n, $r_n > 0$. Since $K = \bigcup_{|w|=n} K_w$, $\mathbb{P}\left[\mathscr{P}_1^t\right] > 0$ for all $t > 0$, regardless of the phase. Recall also the meaning of $\mathbb{P}\langle A \circledast B \rangle$ from Definition 5.5.3.

LOWER SUBCRITICAL: Suppose that $\sum \mathcal{S}^n r_n < \infty$ and let $t > 0$. Then the total rate of M is finite, so $M_{(0,t]}$ is almost surely a finite set. By Lemma 5.5.1, $\mathscr{A}_t$ is almost surely finite.

Let $W(1) = \{w \in W_* \,;\, K_w \subseteq K_1\}$. Since the total rate of M is finite, the rates of events occurring inside K_1 is finite. Recall that there is some $n \in \mathbb{N}$ such that $r_n > 0$ and without loss of generality suppose that $r_1 > 0$, so as

$$\mathbb{P}\left[M_{(0,t]\times\{\emptyset\}} = \emptyset, M_{(0,t]\times W(1)} = \emptyset, M_{(0,t]\times\{2\}} \neq \emptyset\right] > 0.$$

Thus, with positive probability we have $K_1 \subseteq \mathcal{D}_t$ and a non trivial block K_2. Thus $\mathbb{P}\left[\mathscr{P}_2^t\right] > 0$.

Since $\sum \mathcal{S}^n r_n < \infty$, we have also that $\sum r_n < \infty$. By Lemma 7.3.5, if $\mathcal{D}_t \neq 0$ then (a.s.) $\lambda_t^*(\mathcal{D}_t) > 0$. Since $\mathscr{A}_t$ is almost surely finite, $\mathbb{P}\langle \mathscr{P}_1^t \circledast \mathscr{P}_2^t \rangle = 1$.

UPPER SUBCRITICAL: Suppose $\sum \mathcal{S}^n r_n = \infty$ and $\sum r_n < \infty$. By Lemma 7.1.1, $\lambda_t(\mathcal{D}_t) > 0$ with positive probability. If $\lambda_t(\mathcal{D}_t) > 0$ and A_t was finite then $\mathcal{D}_t \neq 0$ and by Lemma 7.1.2 it would follow that $\mathcal{D}_t$ contained a subcomplex, contradicting Lemma 7.1.3. Hence, with positive probability we have that $\lambda_t(\mathcal{D}_t) > 0$ and A_t is infinite. Hence $\mathbb{P}\left[\mathscr{P}_3^t\right] > 0$.

If $\mathcal{D}_t = \emptyset$ then by Lemma 7.3.3, A_t is finite. If $\mathcal{D}_t \neq \emptyset$ then Lemma 7.3.5 shows that (a.s.) $\lambda_t(\mathcal{D}_t) > 0$. By the same reasoning as in the paragraph above, in this case $\mathscr{A}_t$ is infinite. Hence $\mathbb{P}\langle \mathscr{P}_1^t \circledast \mathscr{P}_3^t \rangle = 1$.

SEMICRITICAL: Suppose that $\sum r_n = \infty$, and that $\limsup_n \frac{1}{n}\sum_1^n r_j = 0$. Then, by Lemma 7.4.3, for all $t > 0$ we have $\inf_n m_n^t > 0$. Hence, by Lemma 7.4.2, for all $t > 0$ we have $g^t < \infty$.

By Lemma 7.4.1, $\mathbb{P}\left[\exists n \geq 0, B_n^t = 0\right] < 1$, so by Lemma 7.3.4 we have $\mathbb{P}\left[\mathcal{D}_t \neq \emptyset\right] > 0$. If $\mathcal{D}_t \neq \emptyset$ and A_t is finite then by Lemma 7.1.2 we must have $\lambda_t(\mathcal{D}_t) > 0$, but this does not occur almost surely by Lemma 7.1.1. Hence,

$$\mathbb{P}\left[\mathcal{D}_t \neq \emptyset, A_t \text{ is countably infinite}\right] > 0.$$

By Lemma 7.1.1, $\mathbb{P}\left[\mathcal{D}_t = 0\right] = 1$, and hence we have $\mathbb{P}\left[\mathscr{P}_4^t\right] > 0$.

By Lemma 7.3.3, the set A_t is almost surely finite on the event $\mathcal{D}_t = \emptyset$. If $\mathcal{D}_t \neq \emptyset$ then the same argument as in the above paragraph proves that $\lambda_t(\mathcal{D}_t) = 0$ and also that A_t is countably infinite. Hence $\mathbb{P}\langle \mathscr{P}_1^t \circledast \mathscr{P}_4^t \rangle = 1$.

SUPERCRITICAL: Suppose that $\limsup_n \frac{1}{n}\sum_1^n r_j = \infty$. By Lemma 7.4.3, for all $t > 0$ we have $\inf_n m_n^t = 0$. By Lemma 7.4.1, $\mathbb{P}\left[\exists n \geq 0, B_n^t = 0\right] = 1$; it follows immediately that $\mathcal{D}_t = \emptyset$ by Lemma 7.3.4 and that A_t is finite by 7.3.3. Hence $\mathbb{P}\left[\mathscr{P}_1^t\right] = 1$.

CRITICAL: Suppose that $\limsup_n \frac{1}{n}\sum_1^n r_j \in (0, \infty)$. It follows immediately that $\sum r_n = \infty$. By Lemma 7.4.3, there is some $t_0 \in (0, \infty)$ such that $t \in (0, t_0)$ (where $t_0 \in (0, \infty)$) we have $\inf m_n^t > 0$ and that for $t \in (t_0, \infty)$ we have $\inf m_n^t = 0$.

By Lemma 7.4.2 for $t \in (0, t_0)$ we have also that $g^t < \infty$, and as in the semicritical case we have that $\mathbb{P}\langle \mathscr{P}_1^t \circledast \mathscr{P}_4^t \rangle = 1$. As in the supercritical case, for $t \in (t_0, \infty)$ we have $\mathbb{P}\left[\mathscr{P}_1^t\right] = 1$. ■

Chapter 8

The critical phase

We now turn our attention to the critical phase of our model. Recall that in this phase there is a deterministic $t_0 > 0$, known as the critical time, such that: At time $t < t_0$ there is positive probability of seeing null non-empty dust and an infinity of non-singleton blocks whereas at $t > t_0$ with probability one we see only a finite number of non-singleton blocks and no dust.

8.1 At the critical time

In this section we answer some questions which were left open by Theorem 5.2.1, namely the value of the critical time t_0 and the state of the coalescent at time t_0. We state these results in the following two corollaries to Theorem 5.2.1.

Corollary 8.1.1 *If X is critical then the critical time is given by*

$$t_0 = \frac{\log \mathcal{S}}{\limsup_n \frac{1}{n} \sum_1^n r_j}. \tag{8.1.1}$$

Proof: Let

$$\mathscr{L} = \limsup_n \frac{1}{n} \sum_1^n r_j$$

and let $t^* = \frac{\log \mathcal{S}}{\mathscr{L}}$. Since X is critical we have $0 < \mathscr{L} < \infty$. By Theorem 5.2.1, the critical time t_0 is given by

$$t_0 = \sup\{t > 0 \,;\, \mathbb{P}\left[\mathcal{D}_t \neq \emptyset\right] > 0\}.$$

Consider first when $t < t^*$. Then there exists $\epsilon \in (0, \log \mathcal{S})$ such that $t \leq \frac{\log \mathcal{S} - \epsilon}{\mathscr{L}}$. Hence,

$$m_n^t \geq \left(\exp\left(\log \mathcal{S} - \frac{\log \mathcal{S} - \epsilon}{\mathscr{L}} \frac{1}{n} \sum_1^n r_j \right) \right)^n.$$

By definition of $\mathcal{L}$, there exists $N \in \mathbb{N}$ such that for all $n > N$,

$$\frac{1}{n}\sum_{1}^{n} r_j \leq \frac{\log \mathcal{S} - \frac{\epsilon}{2}}{\log \mathcal{S} - \epsilon}\mathscr{L}.$$

Hence, for all $n > N$,

$$m_n^t \geq \left(\exp\left(\log \mathcal{S} - (\log \mathcal{S} - \epsilon)\frac{\log \mathcal{S} - \frac{\epsilon}{2}}{(\log \mathcal{S} - \epsilon)}\right)\right)^n = (\exp(\epsilon/2))^n.$$

Thus $m_n^t \to \infty$ and hence $\inf m_n^t > 0$. By Lemma 7.4.2, we have $\inf_n m_n^t > 0$ and $g^t < \infty$ for all $t \in (0, t^*)$. By Lemmas 7.3.4 and 7.4.1, $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$ for $t < t^*$.

Now consider $t > t^*$. For some $\epsilon > 0$, $\frac{\log \mathcal{S}+\epsilon}{\mathscr{L}} \leq t$. There exists a subsequence (r_{i_n}) of (r_n) such that for all n,

$$\lim_{n\to\infty} \frac{1}{i_n}\sum_{1}^{i_n} r_j = \limsup_{n\to\infty} \frac{1}{n}\sum_{1}^{n} r_j = \mathscr{L}. \tag{8.1.2}$$

Hence

$$m_{i_n}^t \leq \left(\exp\left(\log \mathcal{S} - \frac{\log \mathcal{S} + \epsilon}{\mathscr{L}}\frac{1}{i_n}\sum_{1}^{i_n} r_j\right)\right)^{i_n}.$$

By (8.1.2), there exists $N \in \mathbb{N}$ such that for all $n > N$,

$$\left|\frac{1}{i_n}\sum_{1}^{i_n} r_j - \mathscr{L}\right| \leq \mathscr{L}\frac{\epsilon/2}{\log \mathcal{S} + \epsilon}.$$

For all such n, $\left|\frac{1}{\mathscr{L}}\frac{1}{i_n}\sum_{1}^{i_n} r_j - 1\right| \leq \frac{\epsilon/2}{\log \mathcal{S}+\epsilon}$, so as

$$\frac{1}{\mathscr{L}}\frac{1}{i_n}\sum_{1}^{i_n} r_j \geq 1 - \frac{\epsilon/2}{\log \mathcal{S} + \epsilon} = \frac{\log \mathcal{S} + \epsilon/2}{\log \mathcal{S} + \epsilon}.$$

Hence, for all $n > N$,

$$m_{i_n}^t \leq \left(\exp\left(\log \mathcal{S} - (\log \mathcal{S} + \epsilon)\frac{\log \mathcal{S} + \epsilon/2}{\log \mathcal{S} + \epsilon}\right)\right)^{i_n} = (\exp(-\epsilon/2))^{i_n}.$$

Thus $m_{i_n}^t \to 0$ and $\inf m_n^t = 0$. By Lemmas 7.3.4 and 7.4.1, $\mathbb{P}[\mathcal{D}_t = \emptyset] = 0$ for $t > t^*$.

Combining the two cases, we have that $t_0 = t^*$, and the proof is complete. ■

Note that setting $\limsup_n \frac{1}{n}\sum_1^n r_j = 0, \infty$ gives $t_0 = \infty, 0$ respectively, which matches the behaviour seen above and below the critical phase.

Corollary 8.1.2 *Suppose X is critical. Then, almost surely, $\mathcal{D}_{t_0} = \emptyset$ and A_{t_0} is finite.*

PROOF: Since X is critical, $\mathscr{L} = \limsup_n \frac{1}{n}\sum_1^n r_j \in (0,\infty)$. By Corollary 8.1.1, $t_0 = \frac{\log \mathcal{S}}{\mathscr{L}}$. Define $a_n \in \mathbb{R}$ by

$$r_n = \mathscr{L} + a_n.$$

Hence

$$\begin{aligned} m_n^{t_0} &= \exp\left(\log \mathcal{S}\left(n - \frac{1}{\mathscr{L}}\sum_1^n r_j\right)\right) \\ &= \exp\left(\log \mathcal{S}\left(n - \frac{1}{\mathscr{L}}\left(n\mathscr{L} + \sum_1^n a_j\right)\right)\right) \\ &= \exp\left(-t_0 \sum_1^n a_j\right) \end{aligned} \tag{8.1.3}$$

We consider two cases. If $\limsup_n \sum_1^n a_j = \infty$ then from (8.1.3) we have $\inf_n m_n^{t_0} = 0$. The same argument as in Theorem 5.2.1 for the supercritical case shows that $\mathbb{P}\left[\mathscr{P}_1^{t_0}\right] = 1$.

If $\limsup \sum_1^n a_j < \infty$ then $\inf m_n^{t_0} > 0$. By Lemma 7.4.4 there exists $C \in (0,\infty)$ such that

$$g^{t_0} \geq \frac{C}{\inf_n m_n^{t_0}} \sum_1^n e^{-r_{n+1}t_0}.$$

Since $\limsup \frac{1}{n}\sum_1^n r_j \in (0,\infty)$, (r_n) has a subsequence (r_{i_n}) such that $\limsup_n r_{i_n} < \infty$. Hence $g^{t_0} = \infty$. By Lemma 7.4.1, $\mathbb{P}\left[\exists n \in \mathbb{N}, B_n^{t_0} = 0\right] = 1$, and hence $\mathbb{P}\left[\mathcal{D}_{t_0} = \emptyset\right] = 1$. By Lemma 7.3.3, $\mathbb{P}\left[\mathscr{P}_1^{t_0}\right] = 1$. ■

8.2 The case of Galton-Watson processes

In this section we examine our model with $r_n = c \in (0,\infty)$. By Theorem 5.2.1, such cases are critical and by Corollary 8.1.1 we have $t_0 = \frac{\log \mathcal{S}}{c}$. This case is of special interest to us because, by Lemma 6.4.2, the GWVEs (B_n^t) are in fact classical Galton-Watson processes. The initial distribution is a Bernoulli(e^{-tr_0}) random variable, and the offspring distribution of a stage n individual is binomial with $\mathcal{S}$ trials and success probability e^{-tc}.

Define

$$|X_t| = |X_{0,t}(K)|, \tag{8.2.1}$$

which is the number of (singleton and non-singleton) blocks in the Segregated Λ-coalescent at time t. By Corollary 8.1.2, $|X_{t_0}|$ is finite and thus (for the the critical phase in general) there is no asymptotic growth of $|X_t|$ to be seen at t_0+ in the almost sure sense. That said,

in the case $r_n = c$ we can obtain a closed form for $\mathbb{E}[|X_t|]$ from Lemma 6.5.1 and deduce the following.

Lemma 8.2.1 *If $r_n = c \in (0, \infty)$, then $\mathbb{E}[|X_{t_0}|] = \infty$ and*

$$\lim_{t \downarrow t_0} \frac{\mathbb{E}[|X_t|]}{(t - t_0)^{-1}} \to \frac{1}{c}\left(1 - \frac{1}{\mathcal{S}}\right) \tag{8.2.2}$$

PROOF: Recall that $t_0 = \frac{\log \mathcal{S}}{c}$. By Theorem 5.2.1 and Corollary 8.1.2 we have that $\mathcal{D}_t = \emptyset$ and $|A_t| < \infty$ for all $t \geq t_0$. Hence, for such t we have $|X_t| = |A_t|$. From Lemma 6.5.1 we have

$$\mathbb{E}[|A_t|] = 1 + (\mathcal{S} - 1) \sum_{n=0}^{\infty} \mathcal{S}^n e^{-tc(n+1)}.$$

Thus

$$\begin{aligned}
\mathbb{E}[|X_t|] &= 1 + (\mathcal{S} - 1) \sum_{n=0}^{\infty} \mathcal{S}^n e^{-tc(n+1)} \\
&= 1 + (\mathcal{S} - 1) e^{-tc} \sum_{n=0}^{\infty} (\mathcal{S} e^{-tc})^n.
\end{aligned}$$

If $t = t_0$ then $\mathcal{S} e^{-t_0 c} = 1$ and we have $\mathbb{E}[|X_t|] = \infty$. If $t > t_0$ then $\mathcal{S} e^{-ct} < 1$, so

$$\begin{aligned}
\mathbb{E}[|X_t|] &= 1 + \frac{(\mathcal{S} - 1) e^{-ct}}{1 - \mathcal{S} e^{-tc}} \\
&= 1 + \frac{(\mathcal{S} - 1) e^{-ct}}{1 - e^{-c(t - t_0)}} \\
&= 1 + \frac{(\mathcal{S} - 1) e^{-tc}}{1 - (1 - c(t - t_0) + \mathcal{O}[(t - t_0)^2])}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\lim_{t \downarrow t_0} \frac{\mathbb{E}[|X_t|]}{(t - t_0)^{-1}} &= \frac{(\mathcal{S} - 1) e^{-ct_0}}{c} \\
&= \frac{1}{c}\left(1 - \frac{1}{\mathcal{S}}\right).
\end{aligned}$$

which completes the proof. ■

Combining Corollary 8.1.2 and Lemma 8.2.1, we have shown that the almost sure behaviour of $|X_t|$ differs from the behaviour of $\mathbb{E}[|X_t|]$ as $t \downarrow 0$.

8.3 Fractal dust

Corollary 8.1.2 raises the issue of precisely how the dust disappears in the run up to time t_0. Lemma 8.3.2 answers this question in a probabilistic sense and following that we also address the question in a spatial sense.

For $s < t$, define $\mathcal{F}_{s,t} = \sigma(M_{(s,t]})$. For $t \in [0,\infty)$ let $\mathcal{E}^t = \inf\{s > t\,;\, M_{(t,s]\times\{w\}} \neq \emptyset\}$. In words, this is the first time after t at which K_w sees a coalescence event. For each $w \in W_*$, $s \in (0,\infty)$ and $t \in [0,\infty)$ define

$$\begin{aligned} Q_{w,t} &= \{\exists \text{ a sequence } (i_n)_{n\in\mathbb{N}} \subseteq S \text{ such that } \forall m \in \mathbb{N}, \mathcal{E}_{wi_1\ldots i_m} > t\} \\ Q^t_{w,s} &= \{\exists \text{ a sequence } (i_n)_{n\in\mathbb{N}} \subseteq S \text{ such that } \forall m \in \mathbb{N}, \mathcal{E}^t_{wi_1\ldots i_m} > s\} \\ R^t_w &= \Omega \setminus \cup_{s>0}(Q^t_{w,s}). \end{aligned}$$

In the language of percolation, $Q_{w,s}$ is the event that $w \in W_*$ is connected to infinity at time t. The set $Q^t_{w,s}$ is the event that a connection between w and infinity that exists at time t will continue to exist until (at least) time $t+s$. The set R^t_w is the event that any connection between w and infinity which might exist at time t will be instantaneously disconnected immediately after time t.

Lemma 8.3.1 *Let $t \in [0,\infty)$ and $w \in W_*$. Then $\mathbb{P}\left[R^t_w\right]$ is either 0 or 1.*

PROOF: For $s < t$, define $\mathcal{F}_{s,t} = \sigma(M_{(s,t]})$. If $s_1 \leq s_2$ then $Q^t_{w,s_2} \subseteq Q^t_{w,s_1}$. Thus, for all $N \in \mathbb{N}$ we have $R^t_w = \Omega \setminus \cup_{n\geq N} Q^t_{w,1/n}$. Noting that $Q^t_{w,s}$ is $\mathcal{F}_{(s,t+s]}$ measurable, we obtain that R^t_w is $\mathcal{F}_{t,t+\frac{1}{N}}$ measurable for all $N \in \mathbb{N}$. The stated result then follows from the Kolmogorov zero-one law. ■

Lemma 8.3.2 *The function $t \mapsto \phi(t) = \mathbb{P}\left[\forall n \in \mathbb{N}_0, B^t_n \neq 0\right]$ is strictly monotone decreasing over $[0,\infty)$. Further, ϕ is left continuous over $t \in [0,\infty)$. If $\phi(s) > 0$ then ϕ is right continuous on $[0,s)$.*

PROOF: Note that $B^t_n \subseteq B^s_n$ for all $s \leq t$; it follows immediately that $\phi(t)$ is decreasing. The time at which clock $\mathcal{E}_w$ ring has a continuous distribution on $[0,\infty)$, so for all $[a,b] \subseteq [0,\infty)$ there is positive probability of having $\mathcal{E}_w \in [a,b]$. It follows from this that $\phi(t)$ is strictly decreasing.

For continuity, note that $\phi(t) = \lim_n \mathbb{P}\left[B^t_n \neq 0\right]$, which is a decreasing limit as $n \to \infty$. Each $\mathcal{E}_w$ has continuous distribution, so the definition of B^t implies that the function $t \mapsto \mathbb{P}\left[B^t_n \neq 0\right]$ is continuous in t. Thus $\phi(t)$ is upper semicontinuous. Since $\phi(t)$ is also decreasing, $\phi(t)$ is left continuous on $[0,\infty)$.

Let $0 \leq t < s$ and be such that $\phi(s) > 0$. In order to show that ϕ is right continuous at t, we must prove that the event

$$A = \{\forall n \in \mathbb{N}_0, B_n^t \neq 0 \text{ and } \forall u > t \; \exists n \in \mathbb{N}_0, B_n^u = 0\} \tag{8.3.1}$$

has probability zero. Note that

$$\begin{aligned}
(8.3.1) &= Q_{\emptyset,t} \cap (\cap_{u>t}(\Omega \setminus Q_{\emptyset,u})) \\
&= Q_{\emptyset,t} \cap (\cap_{u>0}(\Omega \setminus Q_{\emptyset,t+u})) \\
&= Q_{\emptyset,t} \cap (\Omega \setminus \cup_{u>0} Q_{\emptyset,t+u}) \\
&\subseteq Q_{\emptyset,t} \cap (\Omega \setminus \cup_{u>0} Q^t_{\emptyset,u}) \\
&= Q_{\emptyset,t} \cap R^t_\emptyset. \qquad (8.3.2)
\end{aligned}$$

Suppose that A has positive probability. Then by (8.3.2) we have $\mathbb{P}\left[R^t_\emptyset\right] > 0$, which by Lemma 8.3.1 implies that $\mathbb{P}\left[R^t_\emptyset\right] = 1$. By the time homogeneity of our model this means that also $\mathbb{P}\left[R^0_\emptyset\right] = 1$. Hence $\mathbb{P}\left[Q^0_{\emptyset,u}\right] = \mathbb{P}\left[Q_{\emptyset,u}\right] = 0$ for all $u > 0$, which means that B^u is almost surely degenerate for all $u > 0$, in contradiction to our hypothesis that $\phi(s) > 0$. So in fact $\mathbb{P}[A] = 0$, which completes the proof. ■

By Lemma 7.3.1 we have $\phi(t) = \mathbb{P}[\mathcal{D}_t \neq \emptyset]$. Combining Lemma 8.3.2 with Theorem 5.2.1 we have that $t \mapsto \mathbb{P}[\mathcal{D}_t \neq \emptyset]$ is continuous in all but the supercritical phase, in which $\mathbb{P}[\mathcal{D}_t \neq \emptyset] = \mathbb{1}\{t = 0\}$.

In general it does not seem possible to obtain a closed formula for $\mathbb{P}[\mathcal{D}_t \neq \emptyset]$. This leaves us with the question of whether or not the dust also disappears in a spatial sense as $t \uparrow t_0$. We give a partial answer to this, using Hausdorff dimension, as follows.

Note that our definition of a segregated space does not stipulate anything about the spatial proximity of the complexes. The distance between complexes is encoded in the metric D_K, which affects the dimension of K. However, it is easily seen that the genealogy of the flow X is unaffected when D_K is replaced by any topologically equivalent metric. That said, if K is isometrically embedded into some 'standard' metric space $\mathscr{U}$ ($\mathbb{R}^d$, for example), then D_K is canonically specified by $\mathscr{U}$ and dimension becomes a useful quantity.

Remark 8.3.3 *Our model is not the first example of a population model, or even a Fleming-Viot like process, which is set in a spatial but intrinsically dimensionless context; models*

on the hierarchical group are well studied. See Dawson and Greven (1999), Dawson et al. (2008) and the references therein.

The set $\mathcal{D}_t$ is, quite simply, K with randomly chosen complexes removed. Therefore, $\mathcal{D}_t$ is a random fractal, in fact it belongs to the large class of random fractals which are stochastic generalizations of iterated function systems. Recall that we defined iterated function systems in Example 4.1.4.

IFSs have been generalised in many directions, both deterministically and stochastically, and formulas for the Hausdorff dimension of the corresponding attractors have been obtained in increasing generality. There is a large literature which we will not describe here, and instead refer the reader to Durand (2009), Mörters (2010) and the references therein. Generality sufficient to cope with $\mathcal{D}_t$, at least in terms of Hausdorff dimension, seems to have been reached only recently in Durand (2009). Results concerning the Hausdorff measure of the attractors are rarer, and a result corresponding to our context does not seem to be known.

To achieve compatibility with Durand (2009) we must impose some strict conditions on K. A formal statement of the conditions and result (with explicit formulae) appears in Appendix D. The conditions apply, for example, when K is the $\mathcal{S}$-part Cantor set and the result is as follows.

Let $\dim_{\mathcal{H}}(A)$ denote the Hausdorff dimension of $A \subseteq \mathbb{R}^d$.

- If X is critical then, conditional on $\mathcal{D}_t \neq \emptyset$, $t \mapsto \dim_{\mathcal{H}}(\mathcal{D}_t)$ decreases linearly over $(0, t_0)$, from the initial value $\dim_{\mathcal{H}}(K)$ at time $0+$ down to the value 0 at t_0-.

- If X is semicritical, or (lower or upper) subcritical, then, conditional on $\mathcal{D}_t \neq \emptyset$, $\dim_{\mathcal{H}}(\mathcal{D}_t) = \dim_{\mathcal{H}}(K)$ for all $t > 0$.

When we say 'conditional on $\mathcal{D}_t \neq \emptyset$', we mean to condition at fixed time t on the event $\{\mathcal{D}_t \neq \emptyset\}$. Of course, by Theorem 5.2.1 and Corollary 8.1.2, $\mathcal{D}_t$ is a.s. empty if X is supercritical or if X is critical and $t \geq t_0$.

The linear decrease of $\dim_{\mathcal{H}}(\mathcal{D}_t)$ over $(0, t_0)$ which occurs in critical phase shows that, as well as becoming empty with increasing probability, the dust also disappears in a geometric sense as $t \uparrow t_0$.

Perhaps the most striking consequence of the above result is that, in the semicritical phase, the action of the reproduction events is sufficiently weak that it does not change the dimension of $\mathcal{D}_t$. Clearly though, $\mathcal{D}_t$ (which is a $\dim_{\mathcal{H}}(K)$-null set in this phase) must become smaller in some sense as time passes. A description of the way in which this occurs could be very delicate and we do not attempt it here.

Chapter 9

The Segregated Λ-coalescent coming down from infinity

Recall that

$$|X_t| = |X_{0,t}(K)|$$

is the total number of (singleton and non-singleton) blocks in the Segregated Λ-coalescent at time t. In this chapter we study the behaviour of $|X_t|$ in the supercritical phase. We are interested in rate at which $|X_t| \to \infty$ as $t \downarrow 0$, which is known as the rate of coming down from infinity (CDI) of our model. In practice, this means we would like to find a tractable function π such that

$$\lim_{t \downarrow 0} \frac{|X_t|}{\pi(t)}$$

exists in some sense, preferably almost surely or in L^p ($p \geq 1$). As for the Λ-coalescent, the choice of π must depend on the parameters of the model.

Recall that for the Λ-coalescent (Π_t) which came down from infinity, Berestycki et al. (2010) found a deterministic function $\nu : (0, \infty) \to (0, \infty)$ such that $\frac{\Pi_t}{\nu(t)} \to 1$ as $t \downarrow 0$, both almost surely and in L^p. We are not be able to give such a neat result in this chapter, but our results are consistent with the belief that an equivalent result holds for the supercritical phase of the Segregated Λ-coalescent.

9.1 Coming down from infinity

Definition 9.1.1 *We say that the Segregated Λ-coalescent comes down from infinity at $t > 0$ if $\mathbb{P}[|X_t| < \infty] = 1$.*

Corollary 9.1.2 *X comes down from infinity at $t > 0$ if and only if either X is supercritical or X is critical and $t \geq t_0$. In such cases, almost surely*

$$|X_t| = |A_t| = |\overline{\mathscr{B}^t}|$$

and $\mathcal{D}_t = \emptyset$.

PROOF: This follows from Theorem 5.2.1, Corollary 8.1.2 and Lemma 6.4.4. ■

It is easy to see that $|X_t|$ is decreasing in t. As we commented in Section 8.2, X_{t_0} is finite so there is nothing interesting to study in terms of coming down from infinity in the critical phase at time t_0+. However, in the supercritical phase we have the following result.

Lemma 9.1.3 *If X is supercritical then $|X_t| \to \infty$ almost surely as $t \downarrow 0$.*

PROOF: Let $n \in \mathbb{N}$. By Corollary 9.1.2 we have $|X_t| = |\overline{\mathscr{B}^t}| < \infty$ and $\mathcal{D}_t = \emptyset$. By definition of $\mathscr{B}$ we can only have $|\overline{\mathscr{B}^t}| \leq \mathcal{S}^n$ if $\mathcal{E}_w < t$ for some $w \in W_*$ with $|w| \leq n$. However, there are only finitely many such w and for each of them we have $\mathcal{E}_w > 0$ almost surely. So, almost surely there is some $t' > 0$ such that $\mathcal{E}_w > t'$ for all $w \in W_*$ with $|W| \leq n$. Hence $\liminf_{t \downarrow 0} |X_t| \geq \mathcal{S}^n$ almost surely. Since $n \in \mathbb{N}$ is arbitrary, we have the result. ■

In this chapter we write $\mathbb{E}|X_t|$ instead of $\mathbb{E}\left[|X_t|\right]$ and similarly for $|\mathscr{B}^t|$, $|A_t|$ and so on. Combining Corollary 9.1.2 with Lemma 6.5.1 we have that

$$\mathbb{E}|X_t| = 1 + (\mathcal{S} - 1) \sum_{n=0}^{\infty} \mathcal{S}^n e^{-t \sum_0^n r_j} \tag{9.1.1}$$

if X comes down from infinity at t, with the proviso that (9.1.1) may be infinite. Recall from Theorem 5.2.1 that X is supercritical if and only if $\limsup_n \frac{1}{n} \sum_0^n r_j = \infty$.

Lemma 9.1.4 *Suppose that $\lim_{n \to \infty} \frac{1}{n} \sum_0^n r_j = \infty$. Then for all $t > 0$, $\mathbb{E}|X_t| < \infty$.*

PROOF: Fix $t > 0$ and note that (as in Section 7.4),

$$\mathcal{S}^n e^{-t \sum_0^n r_j} = \exp\left(\left(\log \mathcal{S} - t \frac{1}{n} \sum_0^n r_j \right)^n \right).$$

Since $\lim_n \frac{1}{n} \sum_0^n r_j = \infty$ there exists $N \in \mathbb{N}$ such that for all $n > N$ we have

$$\log \mathcal{S} - t \frac{1}{n} \sum_0^n r_j \leq -1.$$

It follows from the above and (9.1.1) that

$$\mathbb{E}|X_t| \leq 1 + (\mathcal{S}-1)\left(\sum_{n=0}^{N} \mathcal{S}^n e^{-t\sum_0^n r_j} + \sum_{n=N+1}^{\infty} e^{-n}\right) < \infty$$

which completes the proof. ■

In Lemma 8.2.1 we saw that in the case $r_n = c$ we had $|X_{t_0}| < \infty$ a.s. but also $\mathbb{E}|X_{t_0}| = \infty$. It is easily seen from (9.1.1) that a similar phenomenon can occur in the supercritical phase; if we choose (r_n) such that $\frac{1}{n}\sum_0^n r_j$ oscillates between 0 and ∞ as $n \to \infty$ then we have both $\mathbb{E}|X_t| = \infty$ and $|X_t| < \infty$ a.s. for all $t > 0$.

Remark 9.1.5 *Functions of the form $f(z) = \sum_{n=0}^{\infty} a_n e^{-\lambda_n z}$, where $(\lambda_n) \subseteq [0, \infty)$ is monotone increasing to ∞, are called generalized Dirichlet series. An analogous result to the well known theorem of radius of convergence for power series exists: each generalized Dirichlet series has an 'abscissa of convergence' $\sigma_c \in [-, \infty, \infty]$ such that the series $\sum_{n=0}^{\infty} a_n e^{-\lambda_n z}$ converges for $z = x + iy$ $(x, y \in \mathbb{R})$ if $x > \sigma_c$ and diverges if $x < \sigma_c$ (see e.g. Hardy and Riesz 1915).*

Lemma 8.2.1 shows that if $r_n = c$ then the abscissa of convergence of $z \mapsto \sum_{n=0}^{\infty} \mathcal{S}^n e^{-z\sum_0^n r_j}$ is $\frac{\log \mathcal{S}}{c}$, whereas Lemma 9.1.4 shows that it is zero if $\lim_{n\to\infty} \frac{1}{n}\sum_0^n r_j = \infty$.

Another well known result concerning generalized Dirichlet series is that, providing we work on the right of the abscissa of convergence, they can be differentiated term by term. Thus in regions where $\mathbb{E}|X_t| < \infty$ we have

$$\frac{d}{dt}\mathbb{E}|X_t| = (\mathcal{S}-1)\sum_{n=0}^{\infty} \mathcal{S}^n \left(-\sum_{j=0}^{n} r_j\right) \exp\left(-t\sum_{j=0}^{n} r_j\right). \tag{9.1.2}$$

Further differentiation shows that

$$(-1)^m \frac{d^m}{dt^m} E|X_t| \geq 0$$

for all $m \in \mathbb{N}$ and $t > 0$, which means that $\mathbb{E}|X_t|$ is completely monotone. This is interesting to us because the speed $\nu(t)$ at which the Λ-coalescent comes down from infinity is also completely monotone.

9.2 Results on the rate of CDI

In this section we offer two results concerning the rate of CDI, both of which are proved in Chapter 10. Their proofs are outlined in Section 9.4. These results are essentially snapshots

of work in progress and, although interesting in their own right, they do not give a complete picture of the CDI behaviour in the supercritical case.

Each of the two results characterizes the rate of CDI for a subset of supercritical cases in a different way. Recall that X is supercritical if and only if

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{0}^{n} r_j = \infty. \tag{9.2.1}$$

Remark 9.2.1 *Our model is supercritical if* $\frac{1}{n}\sum_0^n r_j \to \infty$ *as* $n \to \infty$*, which means that* $\sum_0^n r_j$ *must grow at least linearly. The results in this section require that (a subsequence of)* (r_n) *increases at least geometrically, which forces geometric growth of* $\sum_0^n r_j$*.*

The first setting we consider is described as follows. Suppose that there exists $C \in (1, \infty)$ and $\sigma \in \mathbb{N}$ such that:

[$\mathscr{L}1$] For all n, $r_{\sigma(n+1)} \geq C r_{\sigma n}$.

[$\mathscr{L}2$] For all $m \leq \sigma n$, $r_m \leq r_{\sigma n}$.

Condition [$\mathscr{L}1$] says that (r_n) must have a subsequence $(r_{\sigma n})_{n\in\mathbb{N}}$ which increases at least geometrically. Condition [$\mathscr{L}2$] ensures that $(r_{\sigma n})$ is the maximal such subsequence and also imposes some regularity on the tail of the sequence (r_n). Note that [$\mathscr{L}1$] implies [$\mathscr{L}2$] if $\sigma = 1$.

Let $\pi : (0, \infty) \to (0, \infty)$ be monotone decreasing and such that

$$\pi\left(\frac{1}{r_{\sigma n}}\right) = \mathcal{S}^{\sigma n}. \tag{9.2.2}$$

Note that, since $(r_{\sigma n})_{n\in\mathbb{N}}$ is an increasing sequence, an infinity of examples of such π can always be found. In this setting we have the following result.

Theorem 9.2.2 *Suppose [$\mathscr{L}1$] and [$\mathscr{L}2$]. Then, almost surely,*

$$0 < \liminf_{t\downarrow 0} \frac{|X_t|}{\pi(t)} \leq \limsup_{t\downarrow 0} \frac{|X_t|}{\pi(t)} < \infty. \tag{9.2.3}$$

Example 9.2.3 *The simplest instance of Theorem 9.2.2 is when we can take* $\sigma = 1$*, in which case [$\mathscr{L}1$] and [$\mathscr{L}2$] reduce to asking that for all* n*,* $r_{n+1} \geq C r_n$*. Two examples of such* (r_n) *and* π *are*

- $r_n = c^n$ *where* $c > 1$*; using* $\pi(t) = t^{-\frac{\log \mathcal{S}}{\log c}}$.
- $r_n = c^{c^n}$ *where* $c > 1$*; using* $\pi(t) = (-\log t)^{\frac{\log \mathcal{S}}{\log c}}$.

Both these examples of π *can be found by using* (9.2.2) *and inverting the natural continuous extension to* $(0, \infty)$ *of the map* $n \mapsto r_n$*; that is* $r(x) = c^x$ *in the first case and* $r(x) = c^{c^x}$ *in the second.*

Example 9.2.4 *Let* $(r_n^{(1)})$ *and* $(r_n^{(2)})$ *be two sequences in* $[0, \infty)$ *such that* $r_n^{(1)} \leq r_n^{(2)}$ *for all* $n \in \mathbb{N}$*. Define* (r_n) *to be the interlaced sequence*

$$(r_n) = (r_1^{(1)}, r_1^{(2)}, r_2^{(1)}, r_2^{(2)}, r_3^{(1)}, r_3^{(2)}, \ldots).$$

Suppose also that $r_{n+1}^{(2)} \geq C r_n^{(2)}$ *for all* n*, so that* $(r_n^{(2)})$ *satisfies [*$\mathscr{L}1$*] and [*$\mathscr{L}2$*]. In this case, Theorem 9.2.2 shows that a function* $\pi(t)$ *satisfying* (9.2.3) *can be found using only the sequence* $(r_n^{(2)})$*. That is, the rate of CDI is dominated by the (behaviour caused by the) larger sequence* $(r_n^{(2)})$*.*

Theorem 9.2.2 describes the large scale behaviour of $|X_t|$ as $t \downarrow 0$, but it leaves open the possibility of oscillations between the lim inf and the lim sup. The relation (9.2.2) also leaves some freedom over the choice of the function π. In view of (9.2.3), it is natural to ask whether it is possible to find a deterministic function π which satisfies the stronger relation

$$\lim_{t \downarrow 0} \frac{|X_t|}{\pi(t)} \text{ exists almost surely.} \tag{9.2.4}$$

We might even hope, by comparison with the Λ-coalescent, to choose π in such a way that this limit was 1. We are not able to exhibit such a π in the almost sure sense but we are able to do so if the requirement for almost sure convergence in (9.2.4) is replaced by L^1 convergence. To do so we use the following conditions, again with some constant $C \in (1, \infty)$.

[$\mathscr{M}1$] For all n, $r_{n+1} \geq C r_n$.

[$\mathscr{M}2$] The limits $\lim_{n \to \infty} \frac{r_n}{r_{n+1}}$ and $\lim_{n \to \infty} \frac{1}{r_n} \sum_1^n r_j$ both exist in $[0, \infty)$.

Condition [$\mathscr{M}1$] is just [$\mathscr{L}1$] with $\sigma = 1$, whilst condition [$\mathscr{M}2$] controls irregularity in the tail of (r_n). Condition [$\mathscr{M}1$] implies both [$\mathscr{L}1$] and [$\mathscr{L}2$], meaning that the following theorem uses a more restrictive hypothesis than Theorem 9.2.2.

Theorem 9.2.5 *Suppose [$\mathscr{M}1$] and [$\mathscr{M}2$]. Then*

$$\frac{|X_t|}{\mathbb{E}|X_t|} \to 1$$

in L^1 *as* $t \downarrow 0$.

The condition [$\mathscr{M}2$] imposes quite strict regularity on the tail of (r_n), but there are still many interesting examples of (r_n) to which Theorem 9.2.5 applies. For example, both $r_n = C^n, C^{C^n}$ from Example 9.2.3 are covered by Theorem 9.2.5.

The conditions of Theorems 9.2.2 and 9.2.5 both imply that $\frac{1}{n}\sum_0^n r_j \to \infty$ as $n \to \infty$, meaning that Lemma 9.1.4 holds in both cases. Therefore, it is natural to ask if it is possible to choose π satisfying (9.2.2) such that

$$\lim_{t\downarrow 0} \frac{\pi(t)}{\mathbb{E}|X_t|} = 1. \tag{9.2.5}$$

Numerical estimates indicate that the πs we used in Example 9.2.3 do not satisfy (9.2.5). It seems that the πs from example 9.2.3 are a good approximation to $\mathbb{E}|X_t|$ for small t but fail to take account of $\mathcal{O}(1)$ oscillations; see (9.2.8) below.

In view of Theorems 9.2.2 and 9.2.5, the behaviour of $\mathbb{E}|X_t|$ as $t \downarrow 0$ is of great interest to us. We have the following result, which leads us to a corollary of Theorem 9.2.2.

Lemma 9.2.6 *Suppose [$\mathscr{L}1$] and [$\mathscr{L}2$]. Then there exists $A, A' \in (0,\infty)$ such that for all $n \in \mathbb{N}$,*

$$A'\mathcal{S}^{\sigma n} \le \mathbb{E}|X_{1/r_{\sigma n}}| \le A\mathcal{S}^{\sigma n}.$$

Corollary 9.2.7 *When [$\mathscr{L}1$] and [$\mathscr{L}2$] are satisfied,*

$$0 < \liminf_{t\downarrow 0} \frac{|X_t|}{\mathbb{E}|X_t|} \le \limsup_{t\downarrow 0} \frac{|X_t|}{\mathbb{E}|X_t|} < \infty. \tag{9.2.6}$$

Proof: [Of Lemma 9.2.6.] Condition [$\mathscr{L}1$] implies that X is supercritical, hence by (9.1.1) we have

$$\mathbb{E}|X_{1/r_{\sigma n}}| \ge \mathcal{S}^{\sigma n} \exp\left(-\frac{1}{r_{\sigma n}}\sum_0^{\sigma n} r_j\right). \tag{9.2.7}$$

Using [$\mathscr{L}1$] and [$\mathscr{L}2$] we have

$$\frac{1}{r_{\sigma n}}\sum_0^{\sigma n} r_j \le \frac{r_0}{r_n} + \sum_{j=0}^{n-1}\sum_{k=1}^{\sigma} \frac{r_{j\sigma+k}}{r_{\sigma n}}$$

$$\leq 1 + \sum_{j=0}^{n-1} \sum_{k=1}^{\sigma} \frac{r_{(j+1)\sigma}}{r_{\sigma n}}$$
$$\leq 1 + \sum_{j=1}^{n} \sigma \frac{1}{C^{n-j}}$$
$$\leq 1 + \sigma \sum_{j=0}^{\infty} \frac{1}{C^j}.$$

Hence, $\sup_n \frac{1}{r_{n\sigma}} \sum_0^{n\sigma} r_j < \infty$. The lower bound now follows from (9.2.7).

We now approach the upper bound. Let $m_\sigma = \{k\sigma\,;\, k \in \mathbb{N}\,;\, k\sigma \leq m\}$. We note

$$E|X_{1/r_{n\sigma}}| \leq 1 + (\mathcal{S}-1)\left(\sum_{m=0}^{n\sigma} \mathcal{S}^m + \sum_{m=n\sigma+1}^{\infty} \mathcal{S}^m \exp\left(-\frac{1}{r_{n\sigma}} \sum_0^m r_j\right)\right)$$
$$\leq 1 + (\mathcal{S}-1)\left(\mathcal{S}^{n\sigma+1} + \mathcal{S}^{n\sigma} \sum_{m=n\sigma+1}^{\infty} \mathcal{S}^{m-n\sigma} \exp\left(-\frac{r_{m\sigma}}{r_{n\sigma}}\right)\right)$$
$$= 1 + (\mathcal{S}-1)\left(\mathcal{S}^{n\sigma+1} + \mathcal{S}^{n\sigma} \sum_{j=n}^{\infty} \sum_{k=1}^{\sigma} \mathcal{S}^{j\sigma+k-n\sigma} \exp\left(-\frac{r_{j\sigma}}{r_{n\sigma}}\right)\right)$$

Using $[\mathscr{L}1]$ in the above we have

$$E|X_{1/r_{n\sigma}}| \leq 1 + (\mathcal{S}-1)\left(\mathcal{S}^{n\sigma+1} + \mathcal{S}^{n\sigma} \sum_{j=n}^{\infty} \sum_{k=1}^{\sigma} \mathcal{S}^{j\sigma+k-n\sigma} \exp\left(-C^{j-n}\right)\right)$$
$$\leq 1 + (\mathcal{S}-1)\left(\mathcal{S}^{n\sigma+1} + \mathcal{S}^{n\sigma} \sum_{j=n}^{\infty} \sigma \mathcal{S}^{\sigma} \mathcal{S}^{j\sigma-n\sigma} \exp\left(-C^{j-n}\right)\right)$$
$$= 1 + \mathcal{S}^{n\sigma}(\mathcal{S}-1)\mathcal{S}^{\sigma}\left(1 + \sum_{j=0}^{\infty} \sigma (\mathcal{S}^{\sigma})^j \exp\left(-C^j\right)\right).$$

The above equation implies the upper bound and completes the proof. ■

PROOF: [Of Corollary 9.2.7.] By Lemma 9.2.6, when $[\mathscr{L}1]$ and $[\mathscr{L}2]$ are satisfied, for all $\frac{1}{r_{n\sigma}} \leq t \leq \frac{1}{r_{(n-1)\sigma}}$,

$$\frac{1}{A\mathcal{S}^{\sigma}} \leq \frac{\pi(1/r_{(n-1)\sigma})}{\mathbb{E}|X_{1/r_{n\sigma}}|} \leq \frac{\pi(t)}{\mathbb{E}|X_t|} \leq \frac{\pi(1/r_{n\sigma})}{E|X_{1/r_{(n-1)\sigma}}} \leq \frac{\mathcal{S}^{\sigma}}{A'}.$$

Hence,

$$0 < \liminf_{t \downarrow 0} \frac{\pi(t)}{\mathbb{E}|X_t|} \leq \limsup_{t \downarrow 0} \frac{\pi(t)}{\mathbb{E}|X_t|} < \infty. \tag{9.2.8}$$

The stated result now follows from Theorem 9.2.2. ■

9.3 Open questions

Clearly, the biggest open question arising from this thesis is whether the Segregated Λ-coalescent has a rate $\pi(t)$ of coming down from infinity in the almost sure sense. That is, if there exists a deterministic function π such that

$$\lim_{t \downarrow 0} \frac{|X_t|}{\pi(t)} = 1 \text{ almost surely.} \tag{9.3.1}$$

This is likely to be a hard question; we saw in Chapter 2 that the corresponding result for the Λ-coalescent took many years to establish, with contributions made by several authors.

A second (closely related) question is whether Theorem 9.2.5 can be generalized to L^p convergence and to the entire supercritical case, obtaining an L^p analogue of (9.3.1). The generalization to L^p convergence should involve nothing more than more detailed estimates in the same style as Sections 10.3-10.6, but the extension to (r_n) which do not satisfy $[\mathscr{M}1]$ and $[\mathscr{M}2]$ requires some new ideas.

Proving (or disproving) equation (9.3.1) is likely to need a method of comparing the state of the coalescent continuously across time. Note that our use of $\mathscr{B}^t$ in previous chapters has always been for a fixed time $t > 0$. Our proof of Theorem 9.2.2 will involve analysing the countable family $(\mathscr{B}^{1/r_n})_{n \in \mathbb{N}}$ of GWVE trees together, but at no point in this thesis are we able to analyse the full family $(\mathscr{B}^t)_{t>0}$ probabilistically. It is hoped that an adaptation of the martingale method of Berestycki et al. (2012a) will yield results in this direction. Prof. Vlada Limic and I have recently began collaborating on these questions.

Recall that in Lemma 8.2.1 we showed that if $r_n = c$ then $\mathbb{E}|X_t|$ is approximately $\frac{1}{t-t_0}$ as $t \downarrow t_0$. The critical phase is not limited to $r_n = c$ and it would be interesting to understand $\mathbb{E}|X_t|$ (or more generally, the tail behaviour of $|X_t|$) as $t \downarrow t_0$. We could also ask questions about the critical phase in the run up to the critical time, in particular we could examine $\mathbb{E}\left[|X_t| \,\middle|\, |X_t| < \infty\right]$. This is equivalent to looking at the expected number of boundary points of $\mathscr{B}^t$ when $\mathscr{B}^t$ is conditioned to be finite. In the case $r_n = c$ it is possible to explicitly calculate $\mathbb{E}\left[|X_t| \,\middle|\, |X_t| < \infty\right]$ and show that $\mathbb{E}\left[|X_t| \,\middle|\, |X_t| < \infty\right] \asymp \frac{1}{t_0 - t}$ as $t \uparrow t_0$, which is mildly counter intuitive since the (unconditioned) tree $\mathscr{B}^t$ is shrinking as $t \to \infty$. In general explicit calculations are not possible and the behaviour of $\mathbb{E}\left[|X_t| \,\middle|\, |X_t| < \infty\right]$ as $t \uparrow t_0$ may be difficult to understand. On a similar note, it would be interesting to study behaviour of $\mathbb{E}\left[|X_t| \,\middle|\, |X_t| < \infty\right]$ in the semicritical phase.

The Λ-coalescent literature is very well developed and suggests many aspects of the Segregated Λ-coalescent to analyse. One example is the sampling formulae that we mentioned in Section 2.2. Since our model is very tractable, at least in some respects, it might be possible to obtain more explicit sampling formulae than have been found for the Λ-coalescent. Another question that is suggested by the Λ-coalescent literature concerns total length of the coalescent tree. The Λ-coalescent tree was shown to be of infinite length by Berestycki et al. (2010), essentially because Kingman's coalescent gives an infinite length tree and no Λ-coalescent comes down from infinity faster than Kingman's coalescent. By Theorem 9.2.2 our own model can come down from infinity arbitrarily fast, meaning that there is a phase transition (characterised by some as yet unknown condition) between a finite and infinite length coalescent tree. We do not attempt a list of other such possibilities here and we move on to more general suggestions for future work.

We commented in Section 3.7 that one of main motivations for introducing the Segregated Λ-coalescent was that the dual of the SΛFV process did not come down from infinity. It is not known if it is possible to define the SΛFV process without the limitation on the rate of the reproduction that we discussed in Section 3.5. Whilst our own model is one way around these issues, we had to make some sacrifices in order for our own model come down from infinity; namely the tree structure on K which allowed us to define the model in such a way as a large number of small reproduction events did not cause particles to move very far.

It would interesting to develop other coalescents, possibly similar to our own, such that the coalescent operates in a spatial continuum and is able to come down from infinity. Amongst stochastic flows which do not evolve in jumps, it is common to see an uncountable collection of particles instantaneously coalesce into only a finite number of particles. For example, Arratia (1979) proved that this behaviour occurs for the flow which now better known as Brownian web (see e.g. Fontes et al. 2004), Harris (1984) studied the issue for flows of correlated Brownian particles and, more recently, Le Jan and Raimond (2004) demonstrated such behaviour amongst sticky flows. However, it seems that our own model is the only non-trivial example to date of a stochastic flow which both evolves in jumps and comes down from infinity.

Finally, recall that the SΛFV process has multiple individuals at each site of $\mathbb{R}^d$ whereas

the Segregated Λ-coalescent has only has a single individual starting at each point in space. It would be natural to define an equivalent of our model which started with an uncountable infinity of individuals at each point in space and where (at least, notionally) each individual at an affected site chose independently whether or not to participate in the merger caused by a reproduction event. It seems likely that such a process could be defined and analysed using methods similar to our own, but the details of how to do this and the possible behaviour of the coalescent are not clear.

9.4 Outline of the proofs of Theorems 9.2.2 and 9.2.5

As the formula (9.2.2) suggests, the behaviour of $|X_t|$ in Theorems 9.2.2 and 9.2.5 is closely connected to the behaviour of $|X_{1/r_n}|$. In fact, in our proof of Theorem 9.2.2 we first establish the asymptotic behaviour of $|X_{1/r_{\sigma n}}|$ as $n \to \infty$ and then use monotonicity to interpolate between these times and deduce an asymptotic for $|X_t|$.

The lower bound on $|X_{1/r_{\sigma n}}|$ is found by using direct calculations to show that the number of n-complexes in A_{1/r_n} as $n \to \infty$ is bounded below by a constant multiplied by $\mathcal{S}^n$. The upper bound is more complicated and relies on the GWVE trees $\mathscr{B}^{1/r_n}(v)$ for $v \in W_n$, along with the relation

$$|X_t| \leq \sum_{v \in W_n} |X_{0,t}(K_v)| = \sum_{v \in W_n} |\overline{\mathscr{B}^{1/r_n}(v)}|.$$

We use a strong law of large numbers which permits weak dependence (due to Lyons 1988) to analyse the quantities $|\overline{\mathscr{B}^{1/r_{|v|}}(v)}|$.

In order to prove Theorem 9.2.5, we define a sequence (s_n) such that for all n,

$$\frac{1}{s_n} < \frac{1}{r_n} < \frac{1}{s_{n-1}}.$$

The idea is then to consider $t \in (\frac{1}{s_n}, \frac{1}{s_{n-1}}]$ and show that the behaviour of such t reflects the behaviour seen at precisely $t = \frac{1}{r_n}$. The sequence (s_n) must be chosen carefully, but $[\mathscr{M}1]$ and $[\mathscr{M}2]$ turn out to be suitable conditions to facilitate this comparison.

During our proof of Theorem 9.2.5 there is a separation into two cases, differentiated by whether $\alpha = \lim_{n\to\infty} \frac{r_{n+1}}{r_n}$ is finite or infinite. Note that existence of this limit in $(1, \infty]$ is implied by $[\mathscr{M}2]$. When $\alpha = \infty$, it turns out that in the limit as $n \to \infty$, for $t \in (\frac{1}{s_n}, \frac{1}{s_{n-1}}]$ the set A_t is composed only of complexes of levels $n-1, n$ or $n+1$. What has happened

in this case is that $r_n \to \infty$ so fast that, in the limit, each clock $\mathcal{E}_w$ rings very close to time $1/r_{|w|}$. Thus if $u, w \in W_*$ are such that $|u|$ and $|w|$ are both large with $|u| < |w|$, in the limit as $|u| \to \infty$ we can be confident that $\mathcal{E}_w < \mathcal{E}_u$.

In contrast, when $\alpha < \infty$ it turns out that A_t, again for $t \in (\frac{1}{s_n}, \frac{1}{s_{n-1}}]$, contains complexes of all levels as $n \to \infty$. Fortunately, the limiting proportion of $(n+j)$-complexes within A_{1/r_n} decreases as $|j| \to \infty$, which allows us to establish precise limiting results through considering $A_{1/r_n} \cap \{K_v \, ; \, v \in \cup_{|j| \leq M} W_{n+j}\}$ for arbitrary $M \in \mathbb{N}$. It is not immediately obvious from the arguments given in Sections 10.3-10.6 that this is what makes the proof work, since there we use Lemma 6.2.1 and do our calculations with quantities relating to the size of GWVE trees rather than to the levels of their boundary points.

The proofs themselves will be given in the next chapter and we end this section by reiterating our comment that the results which we stated in this chapter are a snapshot of work in progress. Theorems 9.2.2 and 9.2.5, whilst interesting in their own right, should be viewed as prototypes for future work rather than as finished results. Consequently, during their proofs the reader may have the feeling that [$\mathscr{L}$1]-[$\mathscr{M}$2] give us more firepower than is strictly necessary. Whilst it is my belief that this impression is the correct one, the proofs are quite involved and there is a significant amount of work to be done before these assumptions can be weakened.

Chapter 10

Proof of the CDI results

In this chapter we prove Theorems 9.2.2 and 9.2.5. Sections 10.1 and 10.2 cover the proof of Theorem 9.2.2, and the remaining Sections 10.3-10.6 cover the proof of Theorem 9.2.5.

10.1 Almost sure upper bounds

The upper estimates in our proof Theorem 9.2.2 will rely on the quantity $|\overline{\mathscr{B}^{1/r_{|v|}}(v)}|$, which is the number of boundary points of the GWVE tree $\mathscr{B}^{1/r_{|v|}}(v)$. For ease of notation we write

$$\mathcal{L}_v = \left|\overline{\mathscr{B}^{1/r_{|v|}}(v)}\right| \tag{10.1.1}$$

for the remainder of this section. The following result will form the basis of the upper bound claimed in Theorem 9.2.2.

Lemma 10.1.1 *There exists $C' < \infty$ such that*

$$\limsup_{n\to\infty} \frac{1}{\mathcal{S}^n} \sum_{w\in W_n} \mathcal{L}_w \le C' \tag{10.1.2}$$

almost surely.

Assuming Lemma 10.1.1, the upper bound claimed in Theorem 9.2.2 is proved as follows. Note that

$$|X_t| = |X_{0,t}(K)| \le \sum_{w\in W_n} |X_{0,t}(K_w)|. \tag{10.1.3}$$

If there is some $K_v \in A_t$ such that $K_w \subseteq K_v$ then by definition of A_t, $X_{0,t}(K_w) = X_{0,t}(K_v)$ is a single point. Thus, in this case it is immediate that $|X_{0,t}(K_w)| = |\overline{\mathscr{B}^t(w)}| = 1$. On the

other hand, if there is no such K_v then

$$K_w = \biguplus\{K_u\,;\, K_u \in A_t, K_u \subseteq K_w\}.$$

Hence, by definition of A_t we have $|X_{0,t}(K_w)| = |\{u \in A_t\,;\, K_u \subseteq K_w\}|$, which implies that $|X_{0,t}(K_w)| = |\overline{\mathscr{B}^t(w)}|$ in this case too. Choosing $t = 1/r_n$ and $|w| = n$, in both cases we have

$$|X_{1/r_n}(K_w)| = \mathcal{L}_w. \tag{10.1.4}$$

Combining (10.1.3) and (10.1.4) we have

$$|X_{1/r_n}| \leq \sum_{w \in W_n} \mathcal{L}_w.$$

By the above and Lemma 10.1.1,

$$\limsup_{n\to\infty} \frac{|X_{1/r_n}|}{\mathcal{S}^n} \leq C' < \infty$$

almost surely. For all sufficiently large n we then have

$$\frac{|X_t|}{\pi(t)} \leq \frac{|X_{1/r_{\sigma(n+1)}}|}{\pi(1/r_{\sigma n})} = \mathcal{S}^\sigma \frac{|X_{1/r_{\sigma(n+1)}}|}{\mathcal{S}^{\sigma(n+1)}} \frac{\mathcal{S}^{\sigma n}}{\pi(1/r_{\sigma n})} \leq \mathcal{S}^\sigma C'$$

for t such that $\frac{1}{r_{\sigma(n+1)}} \leq t \leq \frac{1}{r_{\sigma n}}$. Hence

$$\limsup_{t\downarrow 0} \frac{|X_t|}{\pi(t)} \leq \mathcal{S}^\sigma C'. \tag{10.1.5}$$

which proves the upper bound claimed in Theorem 9.2.2.

The remainder of Section 10.1 will be concerned with proving Lemma 10.1.1, using an argument based on the strong law of large numbers. The classical strong law of large numbers does not apply to $(\mathcal{L}_w)_{w\in W_*}$ because the $\mathcal{L}_w$ are not independent, but we will show that the following result of Lyons (1988) does apply.

Theorem 10.1.2 *Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of real valued random variables. Suppose that for some $M < \infty$, for all $n \in \mathbb{N}$,*

$$\mathbb{E}\left[|Y_n|^2\right] \leq M. \tag{10.1.6}$$

Suppose further that

$$\sum_{N=1}^{\infty} \frac{1}{N} \left(\mathbb{E}\left[\left| \frac{1}{N} \sum_{n=1}^{N} Y_n \right|^2 \right] \right)^{1/2} < \infty. \tag{10.1.7}$$

Then $\frac{1}{N}\sum_{n=1}^{N} Y_n \to 0$ almost surely.

Remark 10.1.3 *Theorem 10.1.2 is stated as Theorem 6 in Lyons (1988). The statement there has $M = 1$, but the generalization to $M \in (0, \infty)$ is straightforward.*

For the duration of Section 10.1 we enumerate

$$W_* = \{w(n)\,;\, n \in \mathbb{N}\}, \tag{10.1.8}$$

where the enumeration $w(n)$ is chosen so that $m \leq n$ implies $|w(m)| \leq |w(n)|$. We will apply Theorem 10.1.2 to the centred sequence

$$Y_n = \mathcal{L}_{w(n)} - \mathbb{E}\left[\mathcal{L}_{w(n)}\right] \tag{10.1.9}$$

in Lemma 10.1.4 and then deduce Lemma 10.1.1 from Lemma 10.1.4. First we introduce some notation which we use for the remainder of Section 10.1.

For each $m, n \in \mathbb{N}_0$ let

$$p_{i,n} = \exp\left(-\frac{r_i}{r_n}\right). \tag{10.1.10}$$

Note that $p_{i,n}$ is the probability that a level i clock rang after time $1/r_n$. By $[\mathscr{L}1]$ for $m \geq n$ we have

$$p_{m\sigma,n\sigma} \leq \exp\left(-C^{m-n}\right) \tag{10.1.11}$$

We define

$$N_k = \sum_{i=0}^{k} \mathcal{S}^i$$

and note that $N_k \to \infty$ as $k \to \infty$. Since $N_k = \frac{\mathcal{S}^{k+1}-1}{\mathcal{S}-1}$ we have also that

$$k \leq C_1 \log(N_k) + C_2 \tag{10.1.12}$$

for some $C_1, C_2 \in (0, \infty)$ which depend only on $\mathcal{S}$. Note also that since $\mathcal{S} \geq 2$,

$$N_{k-1} \leq \mathcal{S}^k. \tag{10.1.13}$$

Lemma 10.1.4 *It holds that*

$$\frac{1}{N}\sum_{n=1}^{N} Y_n \to 0$$

almost surely.

PROOF: In view of Theorem 10.1.2, we need only prove (10.1.6) and (10.1.7).

We first prove (10.1.6). Choose $J \in \mathbb{N}$, dependent only upon σ and C, such that

$$\exp(-C^J) \leq \frac{1}{\mathcal{S}^\sigma + 1}.$$

Hence, for all $n, m \in \mathbb{N}$ with $m \geq n + J$, by the above and (10.1.11) we have

$$p_{m\sigma,n\sigma} \leq \exp\left(-C^{m-n}\right) \leq \exp\left(-C^{J\sigma}\right) \leq \frac{1}{\mathcal{S}^\sigma + 1}. \tag{10.1.14}$$

By (10.1.14), the GWVE tree $\mathscr{B}^{1/r_n}(v)$ can be contained inside a GWVE tree $\mathcal{T}(v)$ with initial state 1, initial point v, and stage i offspring distribution χ_i given by:

$$\chi_i = \begin{cases} \mathcal{S} & \text{if } i = 0, 1, 2, \ldots, J\sigma - 1 \\ \mathcal{S} & \text{if } i \geq J\sigma \text{ and } \frac{i}{\sigma} \notin \mathbb{N} \\ \text{Binomial}(\mathcal{S}, \frac{1}{\mathcal{S}^\sigma + 1}) & \text{if } i \geq J\sigma \text{ and } \frac{i}{\sigma} \in \mathbb{N} \end{cases}$$

In words, $\mathcal{T}(v)$ has the following description. Starting from the initial point v and working along W_*:

- Each individual in the first $J\sigma$ generations has $\mathcal{S}$ offspring (i.e. their full complement of offspring).
- For the remainder of the tree we repeat the following: a single generation of individuals which each have Binomial$(\mathcal{S}, \frac{1}{\mathcal{S}^\sigma + 1})$ offspring distributions, followed by $\sigma - 1$ generations of individuals which each have $\mathcal{S}$ offspring.

Remark 10.1.5 *With even more notation we could define $\mathcal{T}(v)$ as a subset of W_* using edge values corresponding to χ_i. The stochastic domination of the offspring distributions of $\mathcal{T}(v)$ by those of $\mathscr{B}^{1/r_n}(v)$ implies the existence of a coupling between $\mathcal{T}(v)$ and $\mathscr{B}^{1/r_n}(v)$ which satisfies $\mathcal{T}(v) \subseteq \mathscr{B}^{1/r_n}(v)$.*

In our usual notation, $|\overline{\mathcal{T}(v)}|$ denotes the number of boundary points of $\mathcal{T}(v)$ and $\mathcal{T}_n(v) = \{vw \in \mathcal{T}\,;\, w \in W_*, |w| = n\}$. By (6.3.3) and Remark 10.1.5 we have

$$\mathcal{L}_v = |\mathscr{B}^{1/r_n}(v)| \leq |\mathcal{T}(v)|. \tag{10.1.15}$$

where $|v| = n$. Recall that $Y_n = \mathcal{L}_{w(n)} - \mathbb{E}\left[\mathcal{L}_{w(n)}\right]$. By (10.1.15), in order to prove (10.1.6) it suffices to establish an upper bound on $\mathbb{E}\left[|\overline{\mathcal{T}(v)}|^2\right]$ which is independent of v.

For $n \leq J\sigma$ we have $\mathcal{T}_n(v) = \mathcal{S}^n \leq \mathcal{S}^{J\sigma}$. For $k \in \{0, \ldots, \sigma - 1\}$ and $n \geq J$, using (6.4.2) we have

$$\mathbb{E}\left[|\mathcal{T}_{J\sigma+n\sigma+k}(v)|\right] = \mathcal{S}^{\sigma J}\left(\mathcal{S}^{n\sigma+k-n}\left(\frac{1}{\mathcal{S}^\sigma + 1}\right)^n\right),$$

which implies

$$\mathcal{S}^{\sigma J}\left(\frac{\mathcal{S}^{\sigma-1}}{\mathcal{S}^{\sigma}+1}\right)^{n} \leq \mathbb{E}\left[|\mathcal{T}_{J\sigma+n\sigma+k}(v)|\right] \leq \mathcal{S}^{\sigma(J+1)}\left(\frac{\mathcal{S}^{\sigma}}{\mathcal{S}^{\sigma}+1}\right)^{n}.$$

Note that $\left(\frac{\mathcal{S}^{\sigma-1}}{\mathcal{S}^{\sigma}+1}\right)^{n}$ and $\left(\frac{\mathcal{S}^{\sigma}}{\mathcal{S}^{\sigma}+1}\right)^{n}$ are both monotone decreasing to 0 as $n \to \infty$. Hence, there are $A_1, A_2 \in (0,\infty)$ and $c_1, c_2 \in (0,1)$ such that for all $n \in \mathbb{N}$ and $v \in W_*$,

$$A_1 c_1^n \leq \mathbb{E}\left[|\mathcal{T}_n(v)|\right] \leq A_2 c_2^n. \tag{10.1.16}$$

It follows immediately that for some $Q \in \mathbb{N}$, for all $n > Q$ and all $v \in W_*$, $\mathbb{E}\left[|\mathcal{T}_n(v)|\right] \leq 1$. Since $n \mapsto A_1 c_1^n$ is decreasing, for all $k \leq n$ we have $A_1 c_1^n \leq \mathbb{E}\left[|\mathcal{T}_k(v)|\right]$. Applying these observations and using (6.1.1) we have

$$\begin{aligned}
&\mathbb{E}\left[|\mathcal{T}_n(v)|^2\right] \\
&\quad = \mathbb{E}\left[|\mathcal{T}_n(v)|\right]^2 + \operatorname{var}\left[|\mathcal{T}_n(v)|\right] \\
&\quad \leq (A_2 c_2^n)^2 + \left(\left(\min_{k=0,\dots,n-1} \mathbb{E}\left[\chi_k\right]\right)^{-1} \prod_{k=0}^{n-1} \mathbb{E}\left[\chi_k\right]\right)\left(\max_{k=0,\dots,n-1} \operatorname{var}\left[\chi_k\right]\right) \sum_{l=0}^{n-1} \prod_{k=l+1}^{n-1} \mathbb{E}\left[\chi_k\right] \\
&\quad \leq (A_2 c_2^n)^2 + \frac{1}{A_1 c_1^n}\left(\prod_{k=0}^{n} A_2 c_2^k\right) \mathcal{S}^2 n \left(\prod_{k=0}^{Q} \mathbb{E}\left[|\mathcal{T}_k(v)|\right]\right) \\
&\quad \leq (A_2)^2 (c_2)^{2n} + \mathcal{S}^2 (A_2)^Q \frac{n}{A_1}\left(\frac{A_2 c_2}{c_1}\right)^n (c_2)^{\frac{n^2}{2}}.
\end{aligned} \tag{10.1.17}$$

It follows from (10.1.17) that there exists $A_3 \in (0,\infty)$ and $c_3 \in (0,1)$ such that for all $n \in \mathbb{N}$ and all $v \in W_*$,

$$\mathbb{E}\left[|\mathcal{T}_n(v)|^2\right] \leq A_3 c_3^n. \tag{10.1.18}$$

Using (6.3.2),

$$\begin{aligned}
\mathbb{E}\left[|\overline{\mathcal{T}(v)}|^2\right] &\leq \mathbb{E}\left[\left(1 + (\mathcal{S}-1)\sum_{n=0}^{\infty} |\mathcal{T}_n(v)|\right)^2\right] \\
&\leq 3 + 3(\mathcal{S}-1)^2 \mathbb{E}\left[\left(\sum_{n=0}^{\infty} |\mathcal{T}_n(v)|\right)^2\right] \\
&= 3 + 3(\mathcal{S}-1)^2\left(\sum_{n=0}^{\infty} \mathbb{E}\left[|\mathcal{T}_n(v)|^2\right] + 2\sum_{n=0}^{\infty}\sum_{m=0}^{n-1} \mathbb{E}\left[|\mathcal{T}_n(v)|\,|\mathcal{T}_m(v)|\right]\right).
\end{aligned}$$

Putting and (10.1.18) into the above with the help of the Cauchy-Schwartz inequality,

$$\mathbb{E}\left[|\overline{\mathcal{T}(v)}|^2\right] \leq 3 + 3(\mathcal{S}-1)^2\left(\sum_{n=0}^{\infty} A_3 c_3^n + 2\sum_{n=0}^{\infty}\sum_{m=0}^{n-1} \mathbb{E}\left[|\mathcal{T}_n(v)|^2\right]^{1/2} \mathbb{E}\left[|\mathcal{T}_m(v)|^2\right]^{1/2}\right)$$

$$\leq 3 + 3(\mathcal{S}-1)^2 \left(\sum_{n=0}^{\infty} A_3 c_3^n + 2 \sum_{n=0}^{\infty} (A_3 c_3^n)^{1/2} \sum_{m=0}^{n-1} (A_3 c_3^m)^{1/2} \right)$$
$$\leq 3 + 3(\mathcal{S}-1)^2 \left(\sum_{n=0}^{\infty} A_3 c_3^n + 2 \sum_{n=0}^{\infty} \frac{A_3}{1-\sqrt{c_3}} (\sqrt{c_3})^n \right).$$

In view of (10.1.15), this proves (10.1.6) with

$$M = 3 + 3(\mathcal{S}-1)^2 \left(\sum_{n=0}^{\infty} A_3 c_3^n + 2 \sum_{n=0}^{\infty} \frac{A_3}{1-\sqrt{c_3}} (\sqrt{c_3})^n \right) < \infty.$$

By the Cauchy-Schwarz inequality we can further deduce that

$$\begin{aligned} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right] &\leq \left(\operatorname{var}\left[\mathcal{L}_{w(n)}\right] \operatorname{var}\left[\mathcal{L}_{w(m)}\right]\right)^{1/2} \\ &\leq \left(\mathbb{E}\left[|\overline{\mathcal{T}(w(n))}|^2 \right] \mathbb{E}\left[|\overline{\mathcal{T}(w(m))}|^2 \right] \right)^{1/2} \\ &\leq M. \end{aligned} \tag{10.1.19}$$

We now move on to proving (10.1.7). Fix $N \in \mathbb{N}$ and choose $j \in \mathbb{N}$ such that $N_{j-1} \leq N \leq N_j$. Note that

$$\begin{aligned} \mathbb{E}\left[\left| \frac{1}{N} \sum_{n=1}^{N} Y_n \right|^2 \right] &= \frac{2}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{n} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right] \\ &\leq \frac{2}{N^2} \sum_{n=1}^{N_j} \sum_{m=1}^{n} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right]. \end{aligned} \tag{10.1.20}$$

Recall that $|w(m)| \leq |w(n)|$ since $m \leq n$ in the above. Because of the tree structure of W_*, $\mathcal{L}_{w(n)}$ and $\mathcal{L}_{w(m)}$ are independent unless $w(m)$ is the first $|w(m)|$ letters of $w(n)$. Now, there is some $l \in \mathbb{N}$ such that $N_l < n \leq N_{l+1}$. Then $|w(n)| = l$ and there are precisely l cases of $m \in \{1, \ldots, n\}$ for which $\mathcal{L}_{w(m)}$ is not independent of $\mathcal{L}_{w(n)}$. Combining this observation with (10.1.19) we have

$$\sum_{m=1}^{n} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right] \leq lM$$

and then applying (10.1.12) we obtain

$$\sum_{m=1}^{n} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right] \leq M(C_1 \log(N_l) + C_2).$$

Putting the above into (10.1.20),

$$\frac{2}{N^2} \sum_{n=1}^{N_j} \sum_{m=1}^{n} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right] \leq \frac{2M}{N^2} \sum_{l=0}^{j} \mathcal{S}^l (C_1 \log(N_l) + C_2)$$

$$\leq \frac{2M(C_1 \log(N_j) + C_2)}{N^2} \sum_{l=0}^{j} \mathcal{S}^l. \tag{10.1.21}$$

Since $N_{j-1} \leq N < N_j$ and $N_j = \mathcal{S}N_{j-1} + 1$ we have

$$N_j \leq \mathcal{S}N + 1. \tag{10.1.22}$$

Similarly,

$$\frac{1}{N} \sum_{l=0}^{j} \mathcal{S}^l \leq \frac{N_j}{N_{j-1}} = \frac{\mathcal{S}N_{j-1} + 1}{N_{j-1}} \leq \mathcal{S} + 1. \tag{10.1.23}$$

Putting (10.1.22) and (10.1.23) into (10.1.21), we obtain

$$\frac{2}{N^2} \sum_{n=1}^{N_j} \sum_{m=1}^{n} \operatorname{cov}\left[\mathcal{L}_{w(n)}, \mathcal{L}_{w(m)}\right] \leq 2M(\mathcal{S}+1) \frac{C_1 \log(\mathcal{S}N + 1) + C_2}{N}$$

Combining the above with (10.1.20) we have

$$\sum_{N=1}^{\infty} \frac{1}{N} \left(\mathbb{E}\left[\left| \frac{1}{N} \sum_{n=1}^{N} Y_n \right|^2 \right] \right)^{1/2} \leq \sqrt{2M(\mathcal{S}+1)} \sum_{N=1}^{\infty} \frac{\sqrt{C_1 \log(\mathcal{S}N + 1) + C_2}}{N^{3/2}} < \infty.$$

Thus (10.1.7) holds. ■

Proof: [Of Lemma 10.1.1.] We note that

$$\begin{aligned} \frac{1}{N_j} \sum_{n=1}^{N_j} Y_n &\geq \frac{1}{N_j} \sum_{\{w \in W_* \,;\, |w| = j\}} \mathcal{L}_w - \mathbb{E}\left[\mathcal{L}_w\right] \\ &\geq \frac{1}{\mathcal{S}} \frac{1}{\mathcal{S}^j} \sum_{\{w \in W_* \,;\, |w| = j\}} \mathcal{L}_w - \mathbb{E}\left[\mathcal{L}_w\right]. \end{aligned}$$

In the above, the second line follows from the first by (10.1.13). Hence, by Lemma 10.1.4,

$$\frac{1}{\mathcal{S}^j} \sum_{\{w \in W_* \,;\, |w| = j\}} \mathcal{L}_{w(n)} - \mathbb{E}\left[\mathcal{L}_w\right] \to 0$$

almost surely as $j \to \infty$. Hence, for sufficiently large j we have (almost surely) that

$$\frac{1}{\mathcal{S}^j} \sum_{\{w \in W_* \,;\, |w| = j\}} \mathcal{L}_w \leq 1 + \frac{1}{\mathcal{S}^j} \sum_{\{w \in W_* \,;\, |w| = j\}} \mathbb{E}\left[\mathcal{L}_w\right].$$

By (10.1.19),

$$\mathbb{E}\left[\mathcal{L}_w\right] \leq \mathbb{E}\left[(\mathcal{L}_w)^2\right] \leq M$$

so as in fact

$$\limsup_{j \to \infty} \frac{1}{\mathcal{S}^j} \sum_{\{w \in W_* \,;\, |w| = j\}} \mathcal{L}_w \leq 1 + M.$$

This completes the proof. ■

10.2 Almost sure lower bounds

We now work towards the lower bound claimed in Theorem 9.2.2. We wish to use GWVEs with different parameters to our usual ones and, as in Theorem 6.5.2, we denote this by the addition of a $\widetilde{\cdot}$ to our notation. Define

$$\begin{aligned} \widetilde{r}_0 &= r_0 \\ \widetilde{r}_m &= r_{n\sigma} \quad \text{where } m \in \mathbb{N} \text{ and } (n-1)\sigma < m \leq n\sigma. \end{aligned} \tag{10.2.1}$$

By $[\mathscr{L}2]$, $\widetilde{r}_m \geq r_m$ for all $m \in \mathbb{N}$. Note also that $(\widetilde{r}_m)$ is monotone increasing.

Let $\widetilde{X}$ denote the segregated Λ-coalescent with coagulation rates $(\widetilde{r}_n)$. For $w \in W_*$, let $\widetilde{E}_w$ be the corresponding clock in K_w, with rate $\widetilde{r}_{|w|}$. Let

$$\widetilde{\mathscr{B}}^t_n = \{w \in W_* \,;\, |w| = n \text{ and for all } j = 0, \ldots, n,\ \widetilde{E}_{w_1,\ldots,w_n} > t\}$$

and let $\widetilde{\mathscr{B}}^t = \bigcup_{n\in\mathbb{N}_0} \widetilde{\mathscr{B}}^t_n$. By Lemma 6.4.2 $\widetilde{\mathscr{B}}^t$ is a GWVE tree with initial node $\emptyset$, initial distribution $\mathbb{1}\{\widetilde{E}_\emptyset > t\}$ and edge values given by $\mathcal{G}_{(w,wi)} = \mathbb{1}\{\widetilde{E}_{wi} > t\}$. By Theorem 6.5.2 we can couple X and $\widetilde{X}$ in such a way that for all $t > 0$, $|\widetilde{X}_t| \leq |X_t|$. In particular, since $\widetilde{r}_{n\sigma} = r_{n\sigma}$ we have

$$|\widetilde{X}_{1/\widetilde{r}_{n\sigma}}| \leq |X_{1/r_{n\sigma}}|. \tag{10.2.2}$$

Recall that $\widetilde{X}_t$ denotes the segregated Λ-coalescent with coagulation rates $(\widetilde{r}_n)$. For $i \leq n$ and $0 \leq a, b < \sigma$ and $i\sigma + a \leq n\sigma + b$, by $[\mathscr{L}1]$,

$$\frac{\widetilde{r}_{i\sigma+a}}{\widetilde{r}_{n\sigma+b}} = \frac{\widetilde{r}_{i\sigma}}{\widetilde{r}_{n\sigma}} \leq \frac{1}{C^{n-i}} \leq \frac{C}{(C^{1/\sigma})^{(\sigma n+b)-(\sigma i+a)}}.$$

Hence, for all $i \leq n$ we have

$$\frac{\widetilde{r}_i}{\widetilde{r}_n} \leq \frac{C}{\left(C^{1/\sigma}\right)^{n-i}}.$$

Note that $C^{1/\sigma} > 1$. Setting $A = C$ and $R = C^{1/\sigma}$ we have that for all $i \leq n$,

$$\widetilde{p}_{i,n} = \exp\left(-\frac{\widetilde{r}_i}{\widetilde{r}_n}\right) \geq \exp\left(-\frac{A}{R^{n-i}}\right) \tag{10.2.3}$$

Lemma 10.2.1 *There exists $\delta > 0$ such that $\delta \leq \prod_{i=0}^{n} \widetilde{p}_{i,n} \leq 1$.*

PROOF: Clearly,

$$1 \geq \prod_{i=0}^{n} \widetilde{p}_{i,n} \geq \lim_{n\to\infty} \exp\left(-A\sum_{i=0}^{n} \frac{1}{R^{n-i}}\right) = \lim_{n\to\infty} \exp\left(-A\sum_{i=0}^{n} \frac{1}{R^{i}}\right).$$

Setting $\delta = \exp\left(-A\sum_{i=0}^{\infty} \frac{1}{R^i}\right) > 0$, we have the result. ■

Our proof of the lower bound claimed in Theorem 9.2.2 will use the limit as $n \to \infty$ of the quantity

$$\widetilde{Z}_n = |\widetilde{\mathscr{B}}_n^{1/\widetilde{r}_n}|. \tag{10.2.4}$$

Lemma 10.2.2 *It holds that*

$$\lim_{n\to\infty} \frac{\widetilde{Z}_n}{\mathcal{S}^n \prod_{i=0}^{n} \widetilde{p}_{i,n}} = 1 \tag{10.2.5}$$

almost surely.

By (6.4.2), equation (10.2.5) is equivalent to the statement that $\frac{\widetilde{Z}_n}{\mathbb{E}[\widetilde{Z}_n]} \to 1$. However, what matters to us here is that $\widetilde{Z}_n \approx \mathcal{S}^n$ (see Lemma 10.2.1).

The programme for the remainder of this section is to prove Lemma 10.2.2 and then complete the proof of Theorem 9.2.2.

Remark 10.2.3 *A similar result to Theorem 10.2.2 was proved for (a subfamily of) supercritical GWVEs in Biggins and D'Souza (1992) using a martingale method. The process $(\widetilde{Z}_n)$ is not a GWVE and we employ different methods below.*

In the language of Biggins and D'Souza, Theorem 10.2.2 shows that $(\widetilde{Z}_n)$ has a single rate of growth. As in Biggins and D'Souza, the driving force of this growth is that the mean grows geometrically; $\mathbb{E}[\widetilde{Z}_n] \asymp \mathcal{S}^n$.

The following lemma is a standard result concerning convergence of random variables, after which we will be ready to give the proof of Theorem 10.2.2.

Lemma 10.2.4 *Let $(Y_n)_{n\in\mathbb{N}_0}$ be real valued random variables and let Y also be a real valued random variable. Suppose that for all $\epsilon > 0$,*

$$\sum_{n=0}^{\infty} \mathbb{P}\big[|Y_n - Y| \geq \epsilon\big] < \infty.$$

Then $Y_n \to Y$ almost surely.

PROOF: [Of Theorem 10.2.2.] By (6.4.3) and Lemma 10.2.1,

$$\operatorname{var}\left[\widetilde{Z}_n\right] \leq \frac{1}{\delta}\mathcal{S}^{2n}\sum_{i=0}^{n}\frac{1-\widetilde{p}_{i,n}}{\mathcal{S}^i}.$$

Using (10.2.3) and the inequality $1-e^{-y}\leq y$, it follows that

$$\operatorname{var}\left[\widetilde{Z}_n\right] \leq \frac{1}{\delta}\mathcal{S}^{2n}\sum_{i=0}^{n}\frac{A/R^{n-i}}{\mathcal{S}^i}.$$

Hence, again using Lemma 10.2.1,

$$\operatorname{var}\left[\frac{\widetilde{Z}_n}{\mathcal{S}^n\prod_{i=0}^{n}\widetilde{p}_{i,n}}\right] \leq \frac{A}{\delta^3}\sum_{i=0}^{n}\frac{1}{\mathcal{S}^i R^{n-i}}.$$

Note that $\mathbb{E}[\widetilde{Z}_n] = \mathcal{S}^n\prod_{i=0}^{n}\widetilde{p}_{i,n}$ by (6.4.2). Fix $\epsilon>0$. By Chebychev's inequality we have

$$\begin{aligned}\mathbb{P}\left[\left|\frac{\widetilde{Z}_n}{\mathbb{E}\left[\widetilde{Z}_n\right]}-1\right|\geq\epsilon\right] &\leq \frac{1}{\epsilon^2}\frac{A}{\delta^3}\sum_{i=0}^{n}\frac{1}{\mathcal{S}^i R^{n-i}}\\ &\leq \frac{A}{\epsilon^2\delta^3}\frac{n+1}{(\mathcal{S}\wedge R)^n.}\end{aligned}$$

Since $\min(\mathcal{S},R)>1$,

$$\sum_{n=0}^{\infty}\mathbb{P}\left[\left|\frac{\widetilde{Z}_n}{\mathbb{E}\left[\widetilde{Z}_n\right]}-1\right|\geq\epsilon\right] \leq \frac{A}{\epsilon^2\delta^3}\sum_{n=0}^{\infty}\frac{n+1}{(\mathcal{S}\wedge R)^n}<\infty.$$

By Lemma 10.2.4 we have the stated result. ■

We now establish the lower bound claimed in in Theorem 9.2.2. By (6.3.2), for all $n\in\mathbb{N}$ we have

$$\begin{aligned}|\overline{\widetilde{\mathscr{B}}^{1/\widetilde{r}_n}}| &\geq (\mathcal{S}-1)|\widetilde{\mathscr{B}}^{1/\widetilde{r}_n}|\\ &\geq (\mathcal{S}-1)|\widetilde{\mathscr{B}}_n^t|\\ &= (\mathcal{S}-1)\widetilde{Z}_n.\end{aligned}$$

By the above and Corollary 9.1.2,

$$(\mathcal{S}-1)\widetilde{Z}_{n\sigma} \leq |\widetilde{X}_{1/r_{n\sigma}}|. \tag{10.2.6}$$

From Lemma 10.2.2 we have

$$\lim_{n\to\infty}\frac{\widetilde{Z}_n}{\mathcal{S}^n\prod_{i=0}^{n}\widetilde{p}_{i,n}}=1$$

almost surely. By Lemma 10.2.1, the above equation implies that $\liminf_{n\to\infty} \frac{\widetilde{Z}_n}{\mathcal{S}^n} \geq \delta$. Combining this with (10.2.2) and (10.2.6) we have that

$$\liminf_{n\to\infty} \frac{|X_{1/r_{n\sigma}}|}{\mathcal{S}^{n\sigma}} \geq \delta(\mathcal{S}-1)$$

almost surely. Hence, for all sufficiently large n and $t \in (0,\infty)$ such that $\frac{1}{r_{(n+1)\sigma}} \leq t \leq \frac{1}{r_{n\sigma}}$ we have

$$\frac{X_t}{\pi(t)} \geq \frac{X_{1/r_{n\sigma}}}{\pi(1/r_{(n+1)\sigma})} = \frac{X_{1/r_{n\sigma}}}{\mathcal{S}^{n\sigma}} \frac{\mathcal{S}^{(n+1)\sigma}}{\pi(1/r_{(n+1)\sigma})} \frac{1}{\mathcal{S}^{\sigma}} \geq \frac{\delta(\mathcal{S}-1)}{\mathcal{S}^{\sigma}}.$$

Thus,

$$\liminf_{t\downarrow 0} \frac{|X_t|}{\pi(t)} \geq \frac{\delta(\mathcal{S}-1)}{\mathcal{S}^{\sigma}}.$$

Combining the above equation with (10.1.5) completes the proof of Theorem 9.2.2. ■

10.3 Decomposition for Theorem 9.2.5

The remainder of Chapter 10 is concerned with the proof of Theorem 9.2.5.

Recall that by Corollary 9.1.2 we have $|X_t| = |\overline{\mathscr{B}^t}|$ and by (6.3.2) we have $|\overline{\mathscr{B}^t}| = (\mathcal{S}-1)|\mathscr{B}^t| + 1$. Hence, in order to prove Theorem 9.2.5 it suffices to show that

$$\mathbb{E}\left[\left|\frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|} - 1\right|\right] \to 1 \tag{10.3.1}$$

as $t \downarrow 0$.

By [$\mathscr{M}2$] the limit

$$\lim_{n\to\infty} \frac{r_{n+1}}{r_n} = \alpha \in (1,\infty] \tag{10.3.2}$$

exists (and is greater that 1 by [$\mathscr{M}1$]). We consider two cases, dependent upon whether $\alpha = \infty$ or $\alpha < \infty$. We refer to the case $\alpha = \infty$ as case (1) and the case $\alpha < \infty$ as case (2).

Define

$$s_n = \left(1 - \left(\frac{r_n}{r_{n+1}}\right)^{1/2}\right) r_n + \left(\frac{r_n}{r_{n+1}}\right)^{1/2} r_{n+1}. \tag{10.3.3}$$

For the remainder of Chapter 10, we define

$$n = n(t) \text{ to be the unique } n \in \mathbb{N} \text{ such that } t \in \left(\frac{1}{s_n}, \frac{1}{s_{n-1}}\right]. \tag{10.3.4}$$

We will usually suppress the dependence on t and write n instead of $n(t)$. We will not use the symbol n to mean anything else for the remainder of Chapter 10. Note that taking limits as $t \downarrow 0$ is the same as taking limits as $n \to \infty$. Clearly $\frac{1}{r_n} \in (\frac{1}{s_n}, \frac{1}{s_{n-1}}]$ and for all t,

$$\frac{1}{r_{n+1}} < \frac{1}{s_n} < t \leq \frac{1}{s_{n-1}} < \frac{1}{r_{n-1}}. \tag{10.3.5}$$

Remark 10.3.1 *The role of s_n is to act as an intermediate point in between r_n and r_{n+1}. Then, the fact that $t \in (s_n, s_{n-1}]$ allows us to think of $t \simeq \frac{1}{r_n}$.*

The precise choice of s_n is to achieve the limits (10.3.9) *and* (10.3.10) *below, in particular for case (1). This is needed in order for the behaviour seen at time $t \in (s_n, s_{n-1}]$ to be similar to that seen at precisely $t = \frac{1}{r_n}$.*

Using [$\mathscr{M}2$] and taking limits as $n \to \infty$ in the relation

$$\frac{r_n}{r_{n+1}} \frac{1}{r_n} \sum_0^n r_j + 1 = \frac{1}{r_{n+1}} \sum_0^{n+1} r_j$$

we obtain that

$$\lim_{n\to\infty} \frac{1}{r_n} \sum_0^n r_j = \begin{cases} 1 & \text{in case (1)} \\ \frac{\alpha}{\alpha-1} & \text{in case (2).} \end{cases} \tag{10.3.6}$$

In case (2) we define

$$\gamma = \left(\alpha^{-1} - \alpha^{-3/2} + \alpha^{-1/2}\right)^{-1} \tag{10.3.7}$$

$$\beta = \left(1 - \alpha^{-1/2} + \alpha^{1/2}\right)^{-1}. \tag{10.3.8}$$

and note that since in this case $\alpha \in (1, \infty)$ we have both $\beta, \gamma \in (0, \infty)$. A simple calculation shows that

$$\lim_{n\to\infty} \frac{r_{n+1}}{s_n} = \begin{cases} \infty & \text{in case (1)} \\ \gamma & \text{in case (2)} \end{cases} \tag{10.3.9}$$

$$\lim_{n\to\infty} \frac{r_n}{s_n} = \begin{cases} 0 & \text{in case (1)} \\ \beta & \text{in case (2)} \end{cases} \tag{10.3.10}$$

Remark 10.3.2 *In some sense, case (1) is a limiting example of case (2).*

We record one more limit which we will need later. For each $k \geq 0$, in case (2),

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{s_n} \sum_{l=0}^{n+k} r_l &= \lim_{n\to\infty} \frac{r_n}{s_n} \left(\frac{1}{r_n} \sum_{l=0}^{n} r_l \right) + \frac{r_{n+1}}{s_n} + \frac{r_{n+1}}{s_n} \frac{r_{n+2}}{r_{n+1}} + \ldots + \frac{r_{n+1}}{s_n} \prod_{l=1}^{k} \frac{r_{n+l}}{r_{n+l-1}} \\ &= \frac{\gamma\alpha}{\alpha - 1} + \gamma + \gamma\alpha + \ldots + \gamma\alpha^{k-1}. \end{aligned} \tag{10.3.11}$$

Our estimates on $|X_t|$ are based on the following inequality.

Lemma 10.3.3 *Let $e_j, e', x_i \geq 0$ and $x_i' > 0$ and suppose that $\sum_i x_i, \sum_i x_i' < \infty$. Then*

$$\left| \frac{\sum_i e_i + \sum_i x_i}{e' + \sum_i x_i'} - 1 \right| \leq \frac{e'}{e' + \sum_i x_i'} + \sum_i \frac{e_i}{e' + \sum_i x_i'} + \sum_i \left| \frac{x_i}{x_i'} - 1 \right| \tag{10.3.12}$$

Remark 10.3.4 *The idea of* (10.3.12) *is that* e', e_i *are small and that* x_i *is close to* x_i'.

PROOF: Note that

$$\begin{aligned}\left|\frac{\sum_i e_i + \sum_i x_i}{e' + \sum_i x_i'} - 1\right| &= \left|\frac{(\sum_i e_i + \sum_i x_i - \sum x_i') - e'}{e' + \sum_i x_i'}\right| \\ &\le \left|\frac{e'}{e' + \sum_i x_i'}\right| + \frac{|\sum_i e_i + \sum_i x_i - \sum_i x_i'|}{e' + \sum_i x_i'} \\ &\le \left|\frac{e'}{e' + \sum_i x_i'}\right| + \left|\frac{\sum_i e_i}{e' + \sum_i x_i'}\right| + \frac{|\sum_i x_i - \sum_i x_i'|}{\sum_i x_i'} \\ &= \frac{e'}{e' + \sum_i x_i'} + \frac{\sum_i e_i}{e' + \sum_i x_i'} + \left|\frac{\sum_i x_i}{\sum_i x_i'} - 1\right|. \end{aligned} \tag{10.3.13}$$

For $a_1, a_2 \ge 0$ and $a_1', a_2' > 0$,

$$\begin{aligned}\left|\frac{a_1 + a_2}{a_1' + a_1'} - 1\right| &= \left|\frac{(a_1 - a_1') + (a_2 - a_2') + (a_1' + a_2')}{a_1' + a_2'} - 1\right| \\ &= \left|\frac{a_1 - a_1'}{a_1' + a_2'} + \frac{a_2 - a_2'}{a_1' + a_2'}\right| \\ &\le \frac{|a_1 - a_1'|}{a_1'} + \frac{|a_2 - a_2'|}{a_2'} \\ &= \left|\frac{a_1}{a_1'} - 1\right| + \left|\frac{a_2}{a_2'} - 1\right|. \end{aligned} \tag{10.3.14}$$

Applying (10.3.14) iteratively to the final term of (10.3.13), we obtain (10.3.12). ■

Let $M \in \mathbb{N}$. Clearly

$$|\mathscr{B}^t| = \sum_{k=0}^{\infty} |\mathscr{B}_k^t|.$$

Recall that $n = n(t)$, defined in (10.3.4). In case (1) we use the decomposition

$$|\mathscr{B}^t| = \overbrace{\sum_{j=0}^{n-M-1} |\mathscr{B}_j^t|}^{\mathcal{E}_0^t} + \sum_{j=n-M}^{n} |\mathscr{B}_j^t| + \overbrace{\sum_{j=n+1}^{\infty} |\mathscr{B}_j^t|}^{\mathcal{E}_1^t}. \tag{10.3.15}$$

whereas in case (2) we use the decomposition

$$|\mathscr{B}^t| = \overbrace{\sum_{j=0}^{n-M-1} |\mathscr{B}_j^t|}^{\mathcal{E}_0^t} + \sum_{j=n-M}^{n+M} |\mathscr{B}_j^t| + \overbrace{\sum_{j=n+M+1}^{\infty} |\mathscr{B}_j^t|}^{\mathcal{E}_2^t}. \tag{10.3.16}$$

Note that the term $\mathcal{E}_0^t$ is the same in both cases.

We now apply Lemma 10.3.3. In case (1), with $e_1 = \mathcal{E}_0^t$, $e_2 = \mathcal{E}_1^t$, $e' = \mathbb{E}[e_1 + e_2]$, $x_i = |\mathscr{B}_i^t|$ for $i = n - M, \dots, n$ and $x_i' = \mathbb{E}[x_i]$, we obtain

$$\left|\frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|} - 1\right| \le \frac{\mathbb{E}\left[\mathcal{E}_0^t + \mathcal{E}_1^t\right]}{\mathbb{E}|\mathscr{B}^t|} + \frac{\mathcal{E}_0^t}{\mathbb{E}\,|\mathscr{B}^t|} + \frac{\mathcal{E}_1^t}{\mathbb{E}\,|\mathscr{B}^t|} + \sum_{j=n-M}^{n} \left|\frac{|\mathscr{B}_j^t|}{\mathbb{E}|\mathscr{B}_j^t|} - 1\right| \tag{10.3.17}$$

In case (2), with $e_1 = \mathcal{E}_0^t$, $e_2 = \mathcal{E}_2^t$, $e' = \mathbb{E}[e_1 + e_2]$, $x_i = |\mathscr{B}_i^t|$ for $i = n - M, \ldots, n + M$ and $x_i' = \mathbb{E}[x_i]$, we obtain

$$\left| \frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|} - 1 \right| \leq \frac{\mathbb{E}\left[\mathcal{E}_0^t + \mathcal{E}_2^t\right]}{\mathbb{E}|\mathscr{B}^t|} + \frac{\mathcal{E}_0^t}{\mathbb{E}\,|\mathscr{B}^t|} + \frac{\mathcal{E}_2^t}{\mathbb{E}\,|\mathscr{B}^t|} + \sum_{j=n-M}^{n+M} \left| \frac{|\mathscr{B}_j^t|}{\mathbb{E}|\mathscr{B}_j^t|} - 1 \right| \tag{10.3.18}$$

In Sections 10.4-10.6 we obtain bounds on the right hand sides of (10.3.17) and (10.3.18).

10.4 Error terms

The first three terms of both (10.3.17) and (10.3.18) are error terms. We seek to bound them from above in this section. At this point we advise the reader to recall (10.3.5) and the definition of $n = n(t)$, since both will be used heavily in what follows.

Lemma 10.4.1 *There exists $A \in (0, \infty)$ such that for all $t > 0$, $A\mathcal{S}^n \leq \mathbb{E}|\mathscr{B}^t|$.*

PROOF: By (6.4.2) and the fact that $t \leq 1/s_{n-1}$,

$$\mathbb{E}|\mathscr{B}^t| \geq \mathcal{S}^{n-1} \exp\left(-\frac{1}{s_{n-1}} \sum_0^{n-1} r_j \right) = \mathcal{S}^{n-1} \exp\left(-\frac{r_{n-1}}{s_{n-1}} \frac{1}{r_{n-1}} \sum_0^{n-1} r_j \right).$$

By (10.3.6), $\sup_n \frac{1}{r_n} \sum_0^n r_j < \infty$ and by (10.3.10), $\sup_n \frac{r_n}{s_n} < \infty$. The result follows. ■

Lemma 10.4.2 *In both cases (1) and (2) we have, for all t,*

$$\frac{\mathcal{E}_0^t}{\mathbb{E}\,|\mathscr{B}^t|} \leq \frac{1}{A\mathcal{S}^M}.$$

PROOF: Note that

$$\sum_{j=0}^{n-M} |\mathscr{B}_j^t| \leq \sum_{j=0}^{n-M-1} \mathcal{S}^j \leq \mathcal{S}^{n-M}.$$

Using Lemma 10.4.1 completes the proof. ■

Lemma 10.4.3 *In case (1), for all t,*

$$\lim_{t \downarrow 0} \mathbb{E}\left[\frac{\mathcal{E}_1^t}{\mathbb{E}\,|\mathscr{B}^t|} \right] = 0$$

PROOF: Using Lemma 10.4.1 we have that

$$\frac{\mathcal{E}_1^t}{\mathbb{E}\,|\mathscr{B}^t|} \leq \frac{1}{A\mathcal{S}^n} \sum_{j=n+1}^{\infty} |\mathscr{B}_j^t|.$$

By (6.4.2) we have $\mathbb{E}\left[|\mathscr{B}_j^t|\right] = \mathcal{S}^j \exp\left(-t\sum_{l=0}^{j} r_l\right)$ and we have $t \geq 1/s_n$, so as

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathcal{E}_1^t}{\mathbb{E}\,|\mathscr{B}^t|}\right] &\leq \frac{1}{A\mathcal{S}^n}\sum_{j=1}^{\infty}\mathcal{S}^{n+j}\exp\left(\frac{-1}{s_n}\sum_{l=0}^{n+j} r_l\right)\\
&\leq \frac{1}{A}\sum_{j=1}^{\infty}\mathcal{S}^{j}\exp\left(\frac{-1}{s_n}\sum_{l=n+1}^{n+j} r_l\right)\\
&= \frac{1}{A}\sum_{j=1}^{\infty}\mathcal{S}^{j}\exp\left(-\frac{r_{n+1}}{s_n} - \frac{r_{n+1}}{s_n}\frac{r_{n+2}}{r_{n+1}} - \ldots - \frac{r_{n+1}}{s_n}\prod_{l=2}^{j}\frac{r_{n+l}}{r_{n+l-1}}\right) \qquad (10.4.1)
\end{aligned}$$

We wish to apply dominated convergence to the above sum as $t \downarrow 0$ (or, equivalently, as $n \to \infty$).

By [$\mathscr{M}$2], noting that in case (1) we have $\frac{r_{m+1}}{r_m} \to \infty$ and $\frac{r_{m+1}}{s_m} \to \infty$, let $N \in \mathbb{N}$ be large enough that for all $m \geq N$, $\frac{r_{m+1}}{r_m}, \frac{r_{m+1}}{s_m} \geq 2$. Then for all $t < 1/s_{N-1}$ we have $n(t) \geq N$ and hence

$$\begin{aligned}
(10.4.1) &\leq \frac{1}{A}\sum_{j=2}^{\infty}\exp\left(j\log\mathcal{S} - \sum_{l=1}^{j}2^l\right)\\
&= \frac{1}{A}\sum_{j=2}^{\infty}\exp\left(j\log\mathcal{S} - 2^j\right) < \infty \qquad (10.4.2)
\end{aligned}$$

Similarly, for all $\mathcal{N} \in \mathbb{N}$ there exists some $\widetilde{N} \in \mathbb{N}$ such that for all $m \geq \widetilde{N}$ we have $\frac{r_m}{s_{m-1}}, \frac{r_{m+1}}{r_m} \geq \mathcal{N}$. Without loss of generality we may choose $\widetilde{N} \geq N$. Hence, for each $j \in \mathbb{N}$, if $t \leq 1/s_{\widetilde{N}-1}$ then $n(t) \geq \widetilde{N}$ and we have

$$\mathcal{S}^j\exp\left(-\frac{r_{n+1}}{s_n} - \frac{r_{n+1}}{s_n}\frac{r_{n+2}}{r_{n+1}} - \ldots - \frac{r_{n+1}}{s_n}\prod_{l=2}^{j}\frac{r_{n+l}}{r_{n+l-1}}\right) \leq \exp\left(-1/2 + j\log\mathcal{S} - \sum_{l=1}^{j}\mathcal{N}^l\right).$$

Thus, for each fixed j,

$$\mathcal{S}^j\exp\left(-\frac{r_{n+1}}{s_n} - \frac{r_{n+1}}{s_n}\frac{r_{n+2}}{r_{n+1}} - \ldots - \frac{r_{n+1}}{s_n}\prod_{l=2}^{j}\frac{r_{n+l}}{r_{n+l-1}}\right) \to 0 \text{ as } t \downarrow 0.$$

By the dominated convergence theorem applied to (10.4.1) we have the result (the domination is supplied, for $n \geq N$, by (10.4.2)). ■

Lemma 10.4.4 *In case (2) we have that*

$$\limsup_{t\downarrow 0}\mathbb{E}\left[\frac{\mathcal{E}_2^t}{|\mathscr{B}^t|}\right] \leq \sum_{j=M+1}^{\infty}\exp\left(j\log\mathcal{S} - \frac{\gamma\alpha}{\alpha-1} - \gamma\sum_{l=1}^{j}\alpha^l\right) < \infty$$

PROOF: We proceed similarly to the proof of Lemma 10.4. Using Lemma 10.4.1, equation (6.4.2) and the fact that $t \geq 1/s_n$, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathcal{E}_2^t}{|\mathscr{B}^t|}\right] &\leq \frac{1}{A\mathcal{S}^n} \sum_{j=n+M+1}^{\infty} \mathcal{S}^j \exp\left(-t\sum_0^j r_l\right) \\
&\leq \frac{1}{A} \sum_{j=M+1}^{\infty} \mathcal{S}^j \exp\left(-\frac{1}{s_n}\sum_0^{n+j} r_l\right) \qquad (10.4.3) \\
&\leq \frac{1}{A} \sum_{j=M+1}^{\infty} \mathcal{S}^j \exp\left(\frac{-1}{s_n}\sum_{l=n+1}^{n+j} r_l\right) \\
&= \frac{1}{A} \sum_{j=M+1}^{\infty} \mathcal{S}^j \exp\left(-\frac{r_{n+1}}{s_n} - \frac{r_{n+1}}{s_n}\frac{r_{n+2}}{r_{n+1}} - \ldots - \frac{r_{n+1}}{s_n}\prod_{l=2}^{j}\frac{r_{n+l}}{r_{n+l-1}}\right) \qquad (10.4.4)
\end{aligned}$$

By $[\mathscr{M}1]$ we have $r_{m+1} \geq Cr_m$ and by $[\mathscr{M}2]$ we have $\alpha \geq C > 1$. Since we are in case (2), using $[\mathscr{M}2]$ and (10.3.10) there is some $N \in \mathbb{N}$ such that for all $m \geq N$, both $\frac{r_{m+1}}{r_m} \geq \frac{1+C}{2}$ and $\frac{r_{m+1}}{s_m} \geq \frac{\gamma}{2}$. For $t \leq 1/s_{N-1}$ we have $n(t) \geq N$ and hence

$$\begin{aligned}
(10.4.4) &\leq \frac{1}{A} \sum_{j=M+1}^{\infty} \exp\left(j\log\mathcal{S} - \frac{\gamma}{2}\sum_{l=0}^{j-1}\left(\frac{1+C}{2}\right)^l\right) \\
&\leq \frac{1}{A} \sum_{j=M+1}^{\infty} \exp\left(j\log\mathcal{S} - \frac{\gamma}{2}\left(\frac{1+C}{2}\right)^{j-1}\right) < \infty \qquad (10.4.5)
\end{aligned}$$

For each $j \in \mathbb{N}$, by (10.3.11) we have

$$\mathcal{S}^j \exp\left(-\frac{1}{s_n}\sum_0^{n+j} r_l\right) \to \mathcal{S}^j \exp\left(-\frac{\gamma\alpha}{\alpha-1} - \gamma\sum_{l=0}^{j-1}\alpha^l\right)$$

as $t \downarrow 0$. Applying dominated convergence (here (10.4.5) supplies the domination for $n \geq N$), we have

$$\lim_{t\downarrow 0} \frac{1}{A} \sum_{j=M+1} \mathcal{S}^j \exp\left(\frac{1}{s_n}\sum_{l=0}^{n+j} r_l\right) = \sum_{j=M+1}^{\infty} \mathcal{S}^j \exp\left(-\frac{\gamma\alpha}{\alpha-1} - \gamma\sum_{l=0}^{j-1}\alpha^l\right).$$

The stated result follows from the above and (10.4.3). ■

10.5 Non-error terms

Lemma 10.5.1 *In case (1), for each $k \in \mathbb{N} \cup \{0\}$, it holds that*

$$\mathbb{E}\left[\left(\frac{|\mathscr{B}_{n-k}^t|}{\mathbb{E}|\mathscr{B}_{n-k}^t|} - 1\right)^2\right] \to 0$$

as $t \downarrow 0$.

PROOF: Fix $k \in \mathbb{N} \cup \{0\}$. We must consider only $t \le 1/s_{k-1}$, so as $n-k \ge 0$, but the limit still makes sense. For such t, using (6.4.2) and the fact that $\frac{1}{s_n} < t \le \frac{1}{s_{n-1}}$,

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{|\mathscr{B}^t_{n-k}|}{\mathbb{E}|\mathscr{B}^t_{n-k}|} - 1\right)^2\right] &= \operatorname{var}\left(\frac{|\mathscr{B}^t_{n-k}|}{\mathbb{E}|\mathscr{B}^t_{n-k}|}\right) \\
&= \sum_{j=0}^{n-k} \frac{1 - e^{-tr_j}}{\mathcal{S}^j e^{-t\sum_0^j r_l}} \\
&\le \sum_{j=0}^{n-k} \frac{1 - e^{-\frac{1}{s_{n-1}} r_j}}{\mathcal{S}^j e^{-\frac{1}{s_n}\sum_0^j r_l}} \\
&\le \sum_{j=0}^{n-k} \frac{r_j}{s_{n-1}} \frac{1}{\mathcal{S}^j} e^{\frac{r_n}{s_n}\frac{1}{r_n}\sum_0^n r_l}
\end{aligned}$$

In the above, to deduce the final line we used the inequality $1 - e^{-y} \le y$, which is valid for all $y \in \mathbb{R}$. By (10.3.6) and (10.3.10), $A' = \sup\{\frac{r_n}{s_n}\frac{1}{r_n}\sum_0^n r_l \,;\, n \in \mathbb{N}\} < \infty$. Hence, using $s_{n-1} < r_{n-1}$,

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{|\mathscr{B}^t_{n-k}|}{\mathbb{E}|\mathscr{B}^t_{n-k}|} - 1\right)^2\right] &\le e^{A'} \sum_{j=0}^{n-k} \frac{r_j}{s_{n-1}} \frac{1}{\mathcal{S}^j} \\
&\le e^{A'} \sum_{j=0}^{n} \frac{r_j}{r_{n-1}} \frac{1}{\mathcal{S}^j} \\
&\le e^{A'} \sum_{j=0}^{n} \frac{1}{C^{n-j-1}\mathcal{S}^j} \\
&\le C e^{A'} \frac{n+1}{(C \wedge \mathcal{S})^{n-1}}.
\end{aligned}$$

We used [$\mathscr{M}$1] in the above to get from the second to third lines. The result follows. ■

Lemma 10.5.2 *In case (2), for each $k \in \mathbb{Z}$, it holds that*

$$\mathbb{E}\left[\left(\frac{|\mathscr{B}^t_{n+k}|}{\mathbb{E}|\mathscr{B}^t_{n+k}|} - 1\right)^2\right] \to 0$$

as $t \downarrow 0$.

PROOF: Fix $k \in \mathbb{Z}$. We proceed similar to Lemma 10.5.1. As before, if $k < 0$, we must consider only $t \le 1/s_{-k-1}$, so as $n+k \ge 0$, but the limit still makes sense. For such t, using that $\frac{1}{s_n} < t \le \frac{1}{s_{n-1}}$,

$$\mathbb{E}\left[\left(\frac{|\mathscr{B}^t_{n+k}|}{\mathbb{E}|\mathscr{B}^t_{n+k}|} - 1\right)^2\right] = \operatorname{var}\left(\frac{|\mathscr{B}^t_{n+k}|}{\mathbb{E}|\mathscr{B}^t_{n+k}|}\right)$$

$$= \sum_{j=0}^{n+k} \frac{1-e^{-tr_j}}{\mathcal{S}^j e^{-t\sum_0^j r_l}}$$
$$\leq \sum_{j=0}^{n+k} \frac{1-e^{-\frac{1}{s_{n-1}}r_j}}{\mathcal{S}^j e^{-\frac{1}{s_n}\sum_0^j r_l}}$$

Using $1-e^{-y} \leq y$ we have

$$\mathbb{E}\left[\left(\frac{|\mathscr{B}^t_{n+k}|}{\mathbb{E}|\mathscr{B}^t_{n+k}|}-1\right)^2\right] = \sum_{j=0}^{n+k} \frac{r_j}{s_{n-1}}\frac{1}{\mathcal{S}^j}\exp\left(\frac{1}{s_n}\sum_0^j r_l\right)$$
$$\leq \sum_{j=0}^{n+k} \frac{r_j}{s_{n-1}}\frac{1}{\mathcal{S}^j}\exp\left(\frac{1}{s_n}\sum_0^{n+(k\vee 0)} r_l\right) \tag{10.5.1}$$

By (10.3.11), $A'' = \sup_n \frac{1}{s_n}\sum_{l=0}^{n+(k\vee 0)} r_l < \infty$. If $k > 0$ then by [$\mathscr{M}$1],

$$A''' = \sup_{j=1,\ldots,k}\sup_{n\geq 0}\frac{r_{n+j}}{r_{n-1}\mathcal{S}^j}$$
$$= \sup_{j=1,\ldots,k}\sup_{n\geq 0}\frac{r_{n+j}}{r_{n+j-1}}\cdots\frac{r_n}{r_{n-1}}\frac{1}{\mathcal{S}^j}$$
$$\leq \left(\sup_{n\in\mathbb{N}}\frac{r_n}{r_{n-1}}\right)^k < \infty.$$

Finiteness in the above follows from [$\mathscr{M}$2]. Applying these two bounds to (10.5.1), along with [$\mathscr{M}$1] the fact that $r_{n-1} < s_{n-1}$, we have

$$\mathbb{E}\left[\left(\frac{|\mathscr{B}^t_{n+k}|}{\mathbb{E}|\mathscr{B}^t_{n+k}|}-1\right)^2\right] \leq e^{A''}\sum_{j=0}^{n+(k\vee 0)}\frac{r_j}{r_{n-1}}\frac{1}{\mathcal{S}^j}$$
$$\leq e^{A''}\left(\sum_{j=0}^{n}\frac{r_j}{r_{n-1}}\frac{1}{\mathcal{S}^j} + \sum_{j=n+1}^{n+(k\vee 0)}\frac{r_j}{r_{n-1}}\frac{1}{\mathcal{S}^j}\right)$$
$$\leq e^{A''}\left(\sum_{j=0}^{n}\frac{1}{C^{n-j-1}\mathcal{S}^j} + \sum_{j=1}^{k\vee 0}\frac{r_{n+j}}{r_{n-1}\mathcal{S}^{n+j}}\right)$$
$$\leq Ce^{A''}\left(\frac{n+1}{(C\wedge\mathcal{S})^{n-1}} + \frac{(k\vee 0)(A''')^{k\vee 0}}{\mathcal{S}^n}\right).$$

Letting $t \to \infty$, which causes $n \to \infty$, the result follows. ■

10.6 Proof of Theorem 9.2.5

Recall that we choose an arbitrary $M \in \mathbb{N}$ in Section 10.3. Let us first address case (1). By Lemmas 10.4.2 and 10.4 bound,

$$\limsup_{t\downarrow 0}\mathbb{E}\left[\frac{\mathbb{E}\left[\mathcal{E}^t_0+\mathcal{E}^t_1\right]}{\mathbb{E}|\mathscr{B}^t|} + \frac{\mathcal{E}^t_0}{\mathbb{E}\,|\mathscr{B}^t|} + \frac{\mathcal{E}^t_1}{\mathbb{E}\,|\mathscr{B}^t|}\right] \leq \frac{2}{A\mathcal{S}^{M-1}}. \tag{10.6.1}$$

By Lemma 10.5.1 (and the fact that L^2 convergence implies L^1 convergence),

$$\lim_{t\downarrow 0}\mathbb{E}\left[\sum_{j=n-M}^{n}\frac{|\mathscr{B}_j^t|}{\mathbb{E}|\mathscr{B}_j^t|}\right]=0. \tag{10.6.2}$$

Putting (10.6.1) and (10.6.2) into (10.3.17),

$$\limsup_{t\downarrow}\mathbb{E}\left|\frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|}-1\right|\leq\frac{2}{A\mathcal{S}^{M-1}}.$$

Since $M\in\mathbb{N}$ is arbitrary, in fact

$$\lim_{t\downarrow 0}\mathbb{E}\left|\frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|}-1\right|=0.$$

Thus we have proved Theorem 9.2.5 in case (1).

We now turn our attention to case (2). Similarly, by Lemmas 10.4.2 and 10.4.4,

$$\begin{aligned}\limsup_{t\downarrow 0}\mathbb{E}&\left[\frac{\mathbb{E}\left[\mathcal{E}_0^t+\mathcal{E}_2^t\right]}{\mathbb{E}|\mathscr{B}^t|}+\frac{\mathcal{E}_0^t}{\mathbb{E}\,|\mathscr{B}^t|}+\frac{\mathcal{E}_2^t}{\mathbb{E}\,|\mathscr{B}^t|}\right]\\&\leq 2\left(\frac{1}{A\mathcal{S}^{M-1}}+\sum_{j=M+1}^{\infty}\exp\left(j\log\mathcal{S}-\frac{\alpha}{\alpha-1}-\sum_{l=1}^{j}\alpha^l\right)\right)\end{aligned} \tag{10.6.3}$$

and by Lemma 10.5.2 (once again, since L^2 convergence implies L^1 convergence),

$$\lim_{t\downarrow 0}\mathbb{E}\left[\sum_{j=n-M}^{n+M}\frac{|\mathscr{B}_j^t|}{\mathbb{E}|\mathscr{B}_j^t|}\right]=0. \tag{10.6.4}$$

Putting (10.6.3) and (10.6.4) into (10.3.18), we obtain

$$\limsup_{t\downarrow}\mathbb{E}\left|\frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|}-1\right|\leq 2\left(\frac{1}{A\mathcal{S}^{M-1}}+\sum_{j=M+1}^{\infty}\exp\left(j\log\mathcal{S}-\frac{\alpha}{\alpha-1}-\sum_{l=1}^{j}\alpha^l\right)\right).$$

Recall that $\alpha>1$, so as

$$\sum_{j=1}^{\infty}\exp\left(j\log\mathcal{S}-\frac{\alpha}{\alpha-1}-\sum_{l=1}^{j}\alpha^l\right)<\infty.$$

Since $M\in\mathbb{N}$ is arbitrary, from (10.6.3) we obtain

$$\lim_{t\downarrow 0}\mathbb{E}\left|\frac{|\mathscr{B}^t|}{\mathbb{E}|\mathscr{B}^t|}-1\right|=0.$$

This proves case (2) and completes the proof of Theorem 9.2.5. ■

Appendix A

A note about Poisson point processes

We refer the reader to Kingman (1993) as a reference for the theory of Poisson Point Processes, which we will not recount here.

It is common to describe several well known jump processes in terms of a suitable Poisson point process. Typically, such a Poisson point process M will have points $(t, y) \in (0, \infty) \times \mathcal{Y}$ (for some measurable space $\mathcal{Y}$) and have rate

$$dt \otimes \beta(dy)$$

for some measure β on $\mathcal{Y}$. When $(t, y) \in M$, the stochastic process will jump at time t, and y will contain information as to precisely what the jump is. Note that Definition 1.2.1 fits this mould.

It will be useful for us to discuss a simple concrete example. Let M have points $(t, y) \in (0, \infty) \times (0, \infty)$ and rate

$$dt \otimes \Pi(dy)$$

where Π is a measure on $(0, \infty)$ satisfying $\int_0^\infty (1 \wedge x)\Pi(dx) < \infty$. The subordinator $(X_t)_{t \geq 0}$ which has no drift and Lévy measure Π is often given the following description. From the initial state $X_0 = 0$,

- Whenever $(t, y) \in M$, the subordinator jumps with size y; that is we set $X_t = X_{t-} + y$.
- In between such jumps, X_t is constant.

This heuristic description of subordinators makes rigorous sense when Π is a finite measure, since then the subordinator jumps at only finite rate. In this case we can simply write

X_t as a sum of finitely many jumps:

$$X_s = \sum_{(t,y)\in M} \mathbb{1}\{s \leq t\} y \tag{A.1}$$

It is well known that subordinators exist for which Π is not a finite measure. In such cases there are infinitely many $(t, y) \in M$ and consequently we say that M has infinite rate. Such subordinators match (at least, as far as our intuition is concerned) the description given above, but in order to make rigorous sense of the situation we need some extra mathematical machinery. In fact, all we need to do is understand the sum in (A.1) as the limit of the increasing sequence of partial sums

$$X_s = \lim_{n\to\infty} \sum_{(t,y)\in M} \mathbb{1}\{s \leq t\}\mathbb{1}\{y > 1/n\}\, y. \tag{A.2}$$

Note that $M \cap [s, t] \times (1/n, \infty)$ is almost surely finite for all $n \in \mathbb{N}$ and $0 < s < t$, so the summation in (A.2) only a finite sum.

The point is that we should take care when defining a stochastic process in terms of an infinite rate Poisson point process; a limiting argument is usually needed to make sense of the heuristic description.

Appendix B

Proof of Theorem 2.3.7

Remark B.1 *I am grateful to an anonymous referee for part of the argument given below for the case* $\mu^{-2} = \infty$.

If $\mu^{-2} < \infty$ then, by Definition 2.1, for each $s < t$, $M \cap \big([0,t] \times (0,1)\big)$ is almost surely finite. Hence, almost surely only finitely many mergers occur before each time $t > 0$, from which it follows that $\mathbb{P}\left[\forall t > 0, N_t^a < \infty\right] = 1$. See also Example 19 of Pitman (1999).

Now suppose that $\mu^{-2} = \infty$. Let $n \in \mathbb{N}$ and let $t > 0$. Let $\epsilon > 0$. Since $\mu^{-1} < \infty$, choose $\delta \in (0,t)$ such that $n(1 - e^{-\delta\mu^{-1}}) < \epsilon$. Since $\mu^{-2} = \infty$,

$$\mathbb{P}\left[M \cap \big((t-\delta,t) \times [\eta,1]\big) \geq n\right] \to 1 \tag{B.1}$$

as $\eta \downarrow 0$. Note that M is time homogeneous so the left hand side of (B.1) does not depend on t. Hence, we can choose $\eta > 0$ such that for all $\delta > 0$, $\mathbb{P}\left[|M \cap \big((t-\delta,t) \times [\eta,1]\big)| \geq n\right] \geq 1 - \epsilon$.

Enumerate

$$M \cap \big((t-\delta,t) \times [\eta,1]\big) = \{(t_i, x_i)\,;\, i = 1, \ldots, K\}$$

and note that $\{K \geq n\} = \{|M \cap (t-\delta,t) \times [\eta,1]| \geq n\}$. By Definition 2.1, each point (x_i, t_i) creates a non-singleton block which we denote by $b_i \subseteq \mathbb{N}$. Also by Definition 2.1, for each i the block b_i is involved in mergers at rate $\mu^{-1} < \infty$. Let B_k be the event that the block b_k is not involved in a merger during $(t_k, t_k + \delta)$ and let $B_k^{\mathbf{c}}$ denote the complementary event. Hence,

$$\begin{aligned}
\mathbb{P}\left[N^a \geq n\right] &\geq \mathbb{P}\left[\{K \geq n\} \cap \big(\cap_{i=1}^{K} B_k\big)\right] \\
&\geq \mathbb{P}\left[K \geq n\right] - \sum_{i=1}^{n} \mathbb{P}\left[B_k^{\mathbf{c}}\right]
\end{aligned}$$

$$\geq 1 - 2\epsilon.$$

Since $\epsilon > 0$ was arbitrary, in fact $\mathbb{P}\left[N_t^a \geq n\right] = 1$. Since $n \in \mathbb{N}$ was arbitrary, in fact

$$\mathbb{P}\left[N_t^a = \infty\right] = 1. \tag{B.2}$$

This establishes the result we are looking for at a single deterministic time $t > 0$.

Recall from Definition 2.1 that the Λ-coalescent is a Markov process. Let $(a_i)_{i\in\mathbb{N}}$ be the (almost surely infinite, by (B.2)) sequence of non-singleton blocks in the Λ-coalescent at time t. Define an equivalence relation $\sim_s$ on $\mathbb{N}$ by

$$i \sim_s j \Leftrightarrow b_i \text{ and } b_j \text{ are in the same block of } \Pi_{t+s}.$$

By the Markov property and Definition 2.1, the process $\widetilde{\Pi}_s = \mathbb{N}/\sim_s$ is a Λ-coalescent with the partition into singletons as its initial state. By Theorem 2.3.5, since $\mu^{-1} < \infty$, $(\widetilde{\Pi}_t)$ does not come down from infinity. Hence, from (B.2) we have $\mathbb{P}\left[\forall s \geq t, |\widetilde{\Pi}_t| = \infty\right] = 1$. Since each block of $\widetilde{\Pi}_s$ corresponds to a unique non-singleton block of Π_{t+s}, this means that $\mathbb{P}\left[\forall s \geq 0, N_{t+s}^a = \infty\right]$. Recall that $t > 0$ was arbitrary, hence $\mathbb{P}\left[\forall t > 0, N_t^a = \infty\right] = 1$. ■

Appendix C

The SΛFV as a Ξ valued process

We give a formal account of some aspects of the discussion in Chapter 3. We recommend Rudin (1991) as a reference for functional analysis.

C.1 The topology on Ξ

We go about placing a topology on the state space Ξ of the SΛFV process. Let

$$C(\mathcal{K}^n) = \{f : \mathcal{K}^n \to \mathbb{R}\,;\, f \text{ is continuous}\},$$

equipped with the supremum norm $||\cdot||_\infty$. Let $\mathcal{M}_F(\mathcal{K})$ denote the space of finite measures on $\mathcal{K}$ and let $\mathscr{P}(\mathcal{K})$ denote the space of probability measures on $\mathcal{K}$. Let

$$\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(\mathcal{K})) = \left\{\rho : \mathbb{R}^d \to \mathcal{M}_F(\mathcal{K})\,;\, \rho \text{ is measurable and } \operatorname{esssup}\left\{||\rho(x)||_{TV}\,;\, x \in \mathbb{R}^d\right\} < \infty\right\}$$

$$\mathcal{L}^1((\mathbb{R}^d)^n, C(\mathcal{K}^n)) = \left\{\Phi : (\mathbb{R}^d)^n \to C(\mathcal{K}^n)\,;\, \Phi \text{ is measurable and } \int_{(\mathbb{R}^d)^n} ||\Phi(z)||_\infty dz < \infty\right\}.$$

Then $\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(\mathcal{K}))$ is a vector space with seminorm

$$||\rho|| = \operatorname{esssup}\left\{||\rho(x)||_{TV}\,;\, x \in \mathbb{R}^d\right\}$$

(here $||\cdot||_{TV}$ denotes the total variation norm) and $\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(\mathcal{K}))$ is a vector space with seminorm

$$||\Phi|| = \int_{(\mathbb{R}^d)^n} ||\Phi(z)||_\infty dz.$$

We set

$$\rho_1 \sim_1 \rho_2 \text{ iff } \{x \in \mathbb{R}^d\,;\, \rho_1(x) \neq \rho_2(x)\} \text{ is Lebesgue null,}$$

and define $L^\infty(\mathbb{R}^d, \mathcal{M}_F(\mathcal{K}))$ to be the quotient of $\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(\mathcal{K}))$ under the equivalence relation $\sim_1$. Then $L^\infty(\mathbb{R}^d, \mathscr{P}(\mathcal{K}))$ is a Banach space. Let $[\rho]_{\sim_1}$ denote the equivalence class of ρ under $\sim_1$.

Similarly, let

$$\Phi_1 \sim_2 \Phi_2 \text{ iff } \{z \in (\mathbb{R}^d)^n \,;\, \Phi_1(z) = \Phi_2(z)\} \text{ is Lebesgue null.}$$

Let $L^1((\mathbb{R}^d)^n, C(\mathcal{K}^n))$ be the quotient of $\mathcal{L}^1((\mathbb{R}^d)^n, C(\mathcal{K}^n))$ by $\sim_2$. Then $L^1((\mathbb{R}^d)^n, C(\mathcal{K}^n))$ is a Banach space equipped with the norm $||\Phi|| = \int_{(\mathbb{R}^d)^n} ||\Phi(z)||_\infty dz$. For ease of notation, for the duration of this section let us write

$$L^\infty = L^\infty(\mathbb{R}^d, \mathcal{M}_F(\mathcal{K}))$$
$$L^1[n] = L^1((\mathbb{R}^d)^n, C(\mathcal{K}^n))$$

and also $L^1 = L^1[1]$.

Proposition C.1 (Evans 1998, Diestel and Uhl 1977) *L^∞ is isometrically isomorphic to a closed subspace of the dual of L^1 via the action*

$$(\rho, \Phi) = \int_{\mathbb{R}^d} \langle \Phi(x), \rho(x) \rangle \, dx,$$

where $\langle \Phi(x), \rho(x) \rangle = \int_{\mathcal{K}} (\Phi(x)(k)) \; (\rho(x)(dk))$.

We induce the weak-star topology on $\Xi \subseteq L^\infty$ from this identification.

Remark C.2 *In general, L^∞ is not isomorphic to the whole dual of L^1. This fails, for example, if $\mathcal{K} = \{0,1\}^{\mathbb{N}}$ with the usual σ-algebra generated by cylinder sets (see the remarks following Definition III.1.3 of Diestel and Uhl 1977).*

Lemma C.3 (Evans 1998) *Ξ is a compact metrizable space.*

Proof: This follows from the separability of L^1 and the Banach-Alaoghu Theorem. ■

If f and g are functions then we write $(f \otimes g)(x,y) = f(x)g(y)$, whereas if $\lambda_1(dx_1)$ and $\lambda_2(dx_2)$ are measure we write $(\lambda_1 \otimes \lambda_2)(d(x_1, x_2)) = \lambda_1(dx_1)\lambda_2(dx_2)$ for the product measure of λ_1 with λ_2. If $\chi : \mathcal{K}^n \to \mathbb{R}$ is measurable and m is a measure on $\mathcal{K}^n$ then we set $\langle \chi, m \rangle = \int_{\mathcal{K}^n} \chi(k) m(dk)$.

Let $n \in \mathbb{N}$, let $\psi : (\mathbb{R}^d)^n \to \mathbb{R}$ be continuous with compact support and for $i = 1, \ldots, n$ let $\chi_i : \mathcal{K} \to \mathbb{R}$ be continuous. We define $I_n : \Xi \to \mathbb{R}$ by

$$I_n(\rho) = I_n\big(\rho \,;\psi, (\chi_i)\big) = \int_{(\mathbb{R}^d)^n} \psi(x_1, \ldots, x_n) \left\langle \bigotimes_{j=1}^n \chi_j \,,\, \bigotimes_{j=1}^n \rho(x_j) \right\rangle dx_1 \ldots dx_n. \qquad \text{(C.1)}$$

Let $C(\Xi)$ denote the set of real valued continuous functions on Ξ. Let $\mathscr{I}$ denote the linear subspace of $C(\Xi)$ which is spanned by the functions I_n (for any n, ψ, χ_i as above) and constant functions.

Lemma C.4 (Evans 1998) *$\mathscr{I}$ is dense in $C(\Xi)$.*

PROOF: Evans proves this without the restriction that ψ has compact support. The modification required to prove our stated result is minimal. ■

Lemma C.5 *$\mathscr{I}$ is a separating, convergence determining class of Ξ.*

PROOF: Evans proves that $\mathscr{I}$ separates points of Ξ. By Lemma 4.3 of Ethier and Kurtz (1986) the compactness of Ξ implies that $M \subseteq C(\Xi)$ is separating iff M is convergence determining. ■

C.2 The generator of the SΛFV process

The set $\mathscr{I}$, which was defined in Appendix C.1, provides suitable set of test functions to define the generator on. The test functions I_n are quite natural from the point of view of biological sampling. The quantity $I_n(\rho)$ tells us the amount of n-tuplets of $\mathbb{R}^d$ (i.e. $(x_j)_{j=1}^n \in (\mathbb{R}^d)^n$) which in state ρ have the genetic types weighted according to χ_j (for the type of the j^{th} individual) and ψ (for locations). With these test functions, in principle we could extract complete information about the distribution of genetic types in any spatial region with positive Lebesgue measure.

Define a linear operator $G : \mathscr{I} \to C(\Xi)$ defined by setting $GI = 0$ if I is a constant function,

$$GI_n\big(\rho\,;\psi(\chi_i)\big) = \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_0^1 \nu_r(du) \int_{\mathcal{K}} \rho(y)(dk) \tag{C.1}$$

$$\int_{dx_1 \ldots dx_n} \psi(x_1, \ldots, x_n) \left(\prod_{\{j\,;\, x_j \notin B_r(y)\}} \big\langle \chi_j, \rho(x_j) \big\rangle \right) \tag{C.2}$$

$$\times \left(\prod_{\{j\,;\, x_j \in B_r(y)\}} \big\langle \chi_j, (1-u)\rho(x_j) + u\delta_k \big\rangle - \prod_{\{j\,;\, x_j \in B_r(y)\}} \big\langle \chi_j, \rho(x_j) \big\rangle \right) \tag{C.3}$$

and extending to $\mathscr{I}$ by linearity. By Lemmas C.4 and C.5, $\mathscr{I}$ is a strong candidate for a suitable class of test functions to use for the generator. Using an adaptation of Evans (1998), Barton et al. (2010a) show that G can be extended to a linear operator $\overline{G}$ on $C(\Xi)$ which is also a Markov generator (in fact, Barton et al. (2010a) consider a two dimensional torus but the generalization to $d \in \mathbb{N}$ is straightforward). Formally, the SΛFV process is defined to be the Ξ-valued Markov process with generator $\overline{G}$.

As usual, we note that the heuristic interpretation of the generators of jump processes connects Definition 3.2.1 to the formula for G. The first three integrals of (C.1) sample a point (y, r, u), which are the event location, event radius and killing proportion. The fourth integral in (C.1) samples the parent type k from $\rho(y)$, whilst (C.2) and (C.3) are the change in $I_n(\rho)$ when a reproduction event with such parameters is applied to ρ.

C.3 Duality of the SΛFV process

Let $B_t = (B_t(m))_{m=1}^n$ be the dual system of Definition 3.4.1 run for time t from initial state B_0. Define an equivalence relation $\sim_t$ on $\{1, \ldots, n\}$ by $n \sim_t m \Leftrightarrow B_{t-}(n) = B_{t-}(m)$. Suppose $\sim_t$ has $l(t)$ equivalence classes and let them be enumerated as $A_t = \{a_1^t, \ldots, a_{l(t)}^t\}$. Note $B_{t-}(a_k^t)$ is well defined for $k \leq l(t)$. For any bounded measurable $F : [0,1]^n \to \mathbb{R}$ and $\rho \in \Xi$ define Υ_n by

$$\Upsilon_n(\rho, B_t; F) = \int_{\mathcal{K}^{l(t)}} F(k_1, \ldots, k_{l(t)}) \bigotimes_{i=1}^{l(t)} \rho(B_t(a_i^t))(dk_i).$$

Let $\mathbb{E}_{\{B_0(m)=x_m\}}$ denote expectation on the probability space of the n-particle dual system of Definition 3.4.1 with initial state $B_0(m) = x_m$ for $m = 1, \ldots, n$. Let $\mathbb{E}_{\rho_0}$ denote expectation on the probability space of the basic SΛFV process ρ_t with initial condition $\rho_0 \in \Xi$.

The following theorem essentially comes from Section 4 of Barton et al. (2010a).

Theorem C.6 *The Markov semigroup of the SΛFV process is characterized by the relation*

$$\mathbb{E}_{\rho_0}\left[I_n(\rho_t, \Phi)\right] = \int_{(\mathbb{R}^d)^n} \mathbb{E}_{\{B_0(m)=x_m\}}\left[\Upsilon_n\big(\rho_0, B_t; \Phi(x_1, \ldots, x_n)\big)\right] dx_1 \ldots dx_n.$$

Appendix D

The Hausdorff dimension of the dust of the Segregated Λ-coalescent

In this section we apply a result of Durand (2009) to deduce the Hausdorff dimension of $\mathcal{D}_t$, when $\mathcal{D}_t \neq \emptyset$. Some of the arguments in Durand (2009) are, like many of those contained in this thesis, heavily based on GWVEs. Durand draws results concerning GWVEs from Lyons (1992).

In order to link our results to those of Durand (2009), we must use some strong assumptions on K. Let $||\cdot||$ denote the Euclidean norm on $\mathbb{R}^d$, and let $\mathcal{L}^d$ denote d dimensional Lebesgue measure. Let A° denote the (topological) interior of the set A, and let the diameter of A be given by $\text{diam}(A) = \sup\{||x-y|| \,;\, x, y \in E\}$. Recall that a similarity f is a function between subsets of $\mathbb{R}^d$ such that for some $\eta \in (0, \infty)$ and all x, y, $||f(x) - f(y)|| = \eta||x-y||$. We write $\eta = \text{lip}(f)$. Recall also that $\dim_{\mathcal{H}}(A)$ denotes the Hausdorff dimension of A (for $A \subseteq \mathbb{R}^d$ this is with respect to the metric $(x, y) \mapsto ||x - y||$).

Definition D.1 *We say that K is D-compatible if $K \subseteq \mathbb{R}^d$ (and $D_K = ||\cdot||$) and:*

1. *For all $w \in W_*$, K_w is compact.*

2. *For all $w \in W_*$ and $i \in S$ there exists a bijective similarity $f^{(w,i)} : K_w \to K_{wi}$.*

3. *There exists $\epsilon, \epsilon' \in (0, 1)$ and a sequence $(l_n) \subseteq [\epsilon, \epsilon']$ such that for all $w \in W$,*

$$\text{lip}(f^{(w,i)}) = l_{|w|}.$$

4. *There exists $\kappa > 0$ such that for all $w \in W_*$, $\mathcal{L}^d(K_w^\circ) \geq \kappa \,\text{diam}(K_w)^d$.*

Note that, in words, $l_{|w|}$ is the contraction ratio of $f^{(w,i)}$, which is the scale factor that

K_w must be shrunk by to get a copy of K_{wi}. Condition *3* says that this depends only on w, and not on i.

Note that condition *1* implies that K is completely segregated. The D-compatibility conditions are unnatural from our point of view, but Hausdorff dimension is often a difficult quantity to calculate and we choose not to discuss to what extent such conditions could be weakened.

The following result is Theorem 1 of Durand (2009), specialized to our situation.

Lemma D.2 *Suppose that K is D-compatible. Fix $t > 0$. For each $s > 0$ and $n \in \mathbb{N}$ let*

$$\alpha_{s,n}^t = \mathcal{S}(l_n)^s e^{-r_{n+1}t}$$

$$\rho^t(s) = \liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \log \alpha_{s,j}$$

Then, if $\mathcal{D}_t \neq \emptyset$,

$$\dim_{\mathcal{H}}(\mathcal{D}_t) = \sup\{s \in [0,\infty)\,;\, \rho^t(s) > 0\}$$

PROOF: It is easy to check that conditions (A)-(C) of Durand (2009) are implied by a combination of $[\mathscr{K}1]$ and the additional assumptions *1-4* on K. The tree-Markov condition (D) of Durand (2009) is also easily checked for the random variables $(\chi_w^t)_{w\in W_*}$. A short calculation (exploiting independence) verifies that our definition of $\alpha_{s,n}^t$ matches equation (9) of Durand (2009). The formula given for $\dim_{\mathcal{H}}(\mathcal{D}_t)$ is then an immediate consequence of Theorem 1 of Durand (2009). ■

We can now describe precisely the behaviour of the Hausdorff dimension of $\mathcal{D}_t$.

Theorem D.3 *Suppose that K is D-compatible and that $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$. Let*

$$\mathscr{L} = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n} r_j \qquad \text{and} \qquad \mathscr{S} = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (-\log l_n).$$

Conditional on $\{\mathcal{D}_t \neq \emptyset\}$,

$$\dim_{\mathcal{H}}(\mathcal{D}_t) = \left(\frac{\log \mathcal{S} - t\mathscr{L}}{\mathscr{S}}\right) \vee 0$$

PROOF: Since X is not supercritical, by Theorem 5.2.1 we have $0 \leq \limsup_n \frac{1}{n}\sum_1^n r_j < \infty$. Note also that by *3* of the D-compatability conditions, $0 \leq -\log \epsilon' \leq -\log(l_n) \leq -\log \epsilon <$

∞. Hence,

$$\begin{aligned}\rho^t(s) &= \liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} (\log \mathcal{S} + s\log(l_j) - t\log(r_j)) \\ &= \liminf_{n\to\infty} \left(\log \mathcal{S} - s\frac{1}{n}\sum_{j=1}^{n}(-\log l_j) - t\frac{1}{n}\sum_{j=1}^{n} r_j \right) \\ &= \log \mathcal{S} - t\limsup_{n\to\infty} \left(\frac{1}{n}\sum_{j=0}^{n} r_j \right) - s\limsup_{n\to\infty} \left(\frac{1}{n}\sum_{j=1}^{n}(-\log l_j) \right).\end{aligned}$$

The result now follows from Lemma D.2. ■

By considering the case $r_n = 0$ in Theorem D.3 (which we usually prefer to ignore because of degeneracy) we can recover the Hausdorff dimension of K, which is given by

$$\dim_{\mathcal{H}}(K) = \frac{\log \mathcal{S}}{\limsup_n \frac{1}{n}\sum_1^n (-\log l_j)}. \tag{D.1}$$

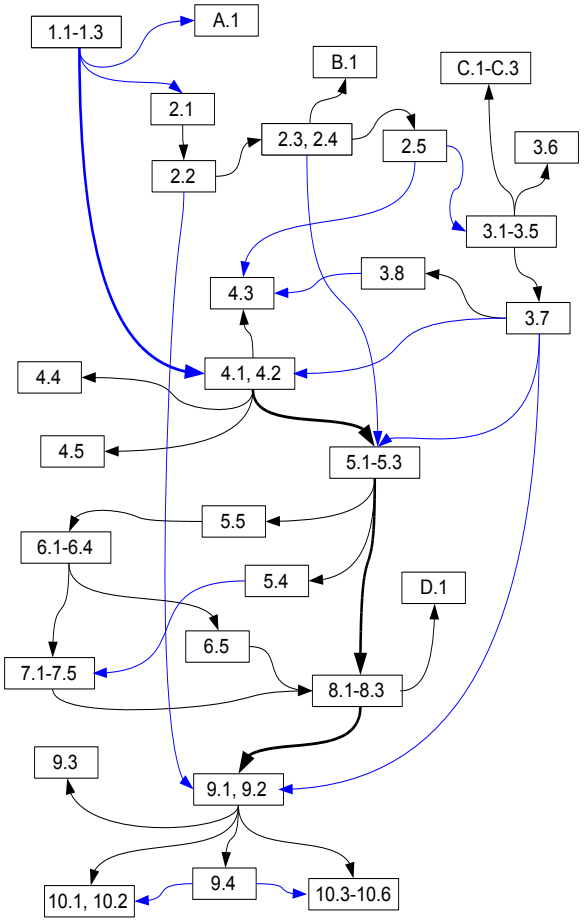
If (l_n) is the constant sequence $l_n = l \in (0,1)$, then $\dim_{\mathcal{H}}(K) = \frac{\log \mathcal{S}}{-\log l}$. This case corresponds to the standard formula for the Hausdorff dimension of a self similar set (see e.g. Falconer 2003). In the case of the $\mathcal{S}$-part Cantor set, which we used in Chapter 4, $l = \frac{1}{2\mathcal{S}-1}$.

In the cases where K is D-compatible, it is possible to use Theorem D.3 to deduce Lemma 7.3.5 and thus slightly shorten our proof of Theorem 5.2.1. Note that Theorem D.3 would not give us the full story about $\mathcal{D}_t$, even when combined with Lemma 7.3.1, because Theorem D.3 does not tell is if $\mathbb{P}[\mathcal{D}_t = \emptyset]$ is in $(0,1)$ or is equal to 1.

Theorem 2 of Durand (2009) gives a recursive formula for what in our notation is $\mathbb{P}\left[\exists n, B_n^t = 0\right]$. A similar recursive formula can be obtained using (7.2.3) and (7.2.4). Following this, Durand (2009) gives relatively complicated conditions which determine if $\mathbb{P}\left[\exists n, B_n^t = 0\right]$ is within $(0,1)$ or equal to one. However, Agresti (1975), Jirina (1976) and Lyons (1992) had already given convenient conditions for degeneracy, at least in so far as our own situation is concerned.

Dependency of sections

The following diagram indicates the major dependencies between different sections of this thesis. A black arrow indicates a mathematical or notational dependency, whereas a blue arrow indicates a dependency in terms of the discussion. The thickened arrows indicate the route recommended to the reader who wishes to do no more than skim over the main results.



Index of notation

See Section 1.3 for general mathematical notation. Conditions [$\mathscr{K}1$]-[$\mathscr{K}5$] can be found in Section 4.1 and conditions [$\mathscr{L}1$]-[$\mathscr{M}2$] can be found in Section 9.2.

Λ, (Π_t)	Λ-coalescent, Definition 2.1.1
$\nu(dx)$	(2.1.2)
μ^n	(2.3.1)
N^s, N^a	(2.3.2), (2.3.3)
ρ_t	Λ-Fleming-Viot process, Definition 2.5.1
$\nu(t)$	rate of CDI for Λ-coalescent, Section 2.2
Ξ	Section 3.1, Appendix C.1
$\rho_t(x), \mu(dr), \nu_r(du)$	SΛFV process, Definition 3.2.1, (3.2.1)
B^a_t, B^n_t, $\sim_t$	Dual of the SΛFV process, Definition 3.4.1
Δ_t	(3.7.3)
S, $\mathcal{S}$	(4.1.1)
W_*, W_n, $\lvert w\rvert$, wi	Section 4.1, just after (4.1.1)
K, D_K, K_w, λ	Definition 4.1.1
$\mathcal{O}$	(4.1.4), Definition 4.1.5
$\mathcal{U}$	just above (4.1.5)
$(r_n), \mathcal{R}, \mathcal{U}_w$	just below (4.1.5)
$M_I, M_{I\times V}$	(4.1.6)
$E^{x,s,t}_m = (u^{x,s,t}_m, w^{x,s,t}_m, p^{x,s,t}_m)$	Definition 4.2.1 (without s,t superscripts), Notation 4.2.4
$\mathcal{C}$	(4.2.1)
$X_{s,t}$	Segregated Λ-coalescent, (4.2.2) and Definition 1.2.1
ξ_t	Segregated Λ-Fleming-Viot process, (4.3.1)
$\mathcal{E}_w$	(5.1.1)
$\mathcal{D}_t$, A_t	Definitions 5.1.1 and 5.1.2
$\mathscr{A}_t$	(5.5.1)

E_*	(6.2.1)
$\mathcal{T}$, $\overline{\mathcal{T}}$, $\mathcal{T}_w$, $\mathcal{G}_e$	GWVE tree, Definition 6.3.1
$\mathscr{B}^t(v)$, $\mathscr{B}^t$, B_n^t	Definition 6.4.1 and just below
$I(\mathscr{B}^t)$, Q, $\mathbb{P}_Q$	(7.2.1) and (7.2.2)
m_n^t, g^t	(7.4.1) and (7.4.2)
t_0,	just below Table 5.2, (8.1.1)
$\lvert X_t\rvert$	(8.2.1)
$\mathcal{F}_{s,t}, \mathcal{E}^t, Q_{w,t}, Q_{w,t}^s, R_w^t$	Section 8.3, just before Lemma 8.3.1
π	(9.2.2)
C, σ	Section 9.2
$\mathcal{L}_v$, Y_n	(10.1.1) and (10.1.9)
$w(n)$	(10.1.8)
$p_{i,n}$	(10.1.10)
$(\widetilde{r}_n)$, $\widetilde{X}$, $\widetilde{\cdot}$	(10.2.1), Section 10.2
δ	Lemma 10.2.1
$\widetilde{Z}_n$	(10.2.4)
$n(t)$	(10.3.4)
α, γ, β	(10.3.2), (10.3.7), (10.3.8)
$\mathcal{E}_0^t$, $\mathcal{E}_1^t$, $\mathcal{E}_2^t$	(10.3.15) and (10.3.16)

Bibliography

A. Agresti. On the extinction times of varying and random environment branching processes. *J. Appl. Probab.*, 12(1):39–46, 1975.

O. Angel, N. Berestycki, and V. Limic. Global divergence of spatial coalescents. *Probab. Th. Rel. Fields*, Published online:1–45, 2010.

R. Arratia. Coalescing Brownian motions on the line. *Ph.D. Thesis, University of Wisconsin, Madison*, 1979.

K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, 1972.

N. H. Barton, A. M. Etheridge, and A. Véber. A new model for evolution in a spatial continuum. *Electron. J. Probab.*, 15:162–216, 2010a.

N. H. Barton, J. Kelleher, and A. M. Etheridge. A new model for large scale population dyanmics: quantifying phylogeography. *Evolution*, 64:2701–2715, 2010b.

N. H. Barton, A. M. Etheridge, and A. Véber. Modelling evolution in a spatial continuum. *J. Stat. Mech. (to appear)*, 2012.

J. Berestycki, N. Berestycki, and J. Schweinsberg. Small-time behavior of β-coalescents. *Ann. Inst. H. Poincare Probab. Statis.*, 44(2):214–238, 2008.

J. Berestycki, N. Berestycki, and V. Limic. The Λ-coalescent speed of coming down from infinity. *Ann. Probab.*, 38:207–233, 2010.

J. Berestycki, N. Berestycki, and V. Limic. A small-time coupling between Λ-coalescents and branching processes. *http://arxiv.org/abs/1101.1875*, 2012a.

J. Berestycki, N. Berestycki, and V. Limic. Asymptotic sampling formulae for Λ-coalescents. *http://arxiv.org/abs/1201.6512*, 2012b.

N. Berestycki. *Recent Progress In Coalescent Theory*. Ensaios Matematicos, vol. 16. 2009.

N. Berestycki, A. M. Etheridge, and M. Hutzenthaler. Survival, extinction and ergodicity in a spatially continuous population model. *Markov Process. Rel. Fields*, 15(3):265–288, 2009.

N. Berestycki, A. M. Etheridge, and A. Véber. Large scale behaviour of the Spatial Λ-Fleming Viot process. *Ann. Inst. H. Poincare Probab. Statis. (to appear)*, 2012c.

J. Bertoin. *Random Fragmentation and Coagulation Processes.* Cambridge University Press, 2006.

J. Bertoin and J. F. Le Gall. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Th. Rel. Fields*, 117:249–266, 2000.

J. Bertoin and J. F. Le Gall. Stochastic flows associated to coalescent processes. *Probab. Th. Rel. Fields*, 126:261–288, 2003.

J. Bertoin and J. F. Le Gall. Stochastic flows associated to coalescent processes III. Limit theorems. *Illinois J. Math.*, 50:147–181, 2006.

J. Biggins and J. D'Souza. The supercritical Galton-Watson process in varying environments. *Stochastic Process. Appl.*, 42:39–47, 1992.

M. Birkner, J. Blath, M. Capaldo, A. M. Etheridge, M. Möhle, J. Schweinsberg, and A. Wakolbinger. α-stable branching and β-coalescents. *Electron. J. Probab.*, 10(9):303–325, 2005.

E. Bolthausen and A. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197:247–276, 1998.

J. D. Church. Composition limit theorems for probability generating functions. *MRC Techn. Summ. Rept.*, (732):1–41, 1967.

D. Dawson, A. Greven, F. den Hollander, R. Sun, and J. Swart. The renormalization transformation for two-type branching models. *Ann. Inst. H. Poincare Probab. Statis.*, 44(6):1038–1077, 2008.

D. A. Dawson and A. Greven. Hierarchically interacting Fleming-Viot processes with selection and mutation: Multiple space time scale analysis and quasi-equilibria. *Electron. J. Probab.*, 4:1–81, 1999.

J. Diestel and J. J. Uhl. *Vector Measures.* Providence, 1977.

P. Donnelly and T. G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.*, 9(4):1091–1148, 1999.

J. D'Souza. The rates of groth of the Galton-Watson process in varying environments. *Adv. Appl. Prob.*, 26:698–714, 1994.

A. Durand. Random fractals and tree-indexed Markov chains. *Rev. Mat. Iberoam.*, 25(3): 1089–1126, 2009.

A. M. Etheridge. Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ.*, 80:121–144, 2008.

A. M. Etheridge. *Some mathematical models from population genetics*, volume 2012 of *Lecture Notes in Mathematics, École d'Été de Probabilits de Saint-Flour XXXIX-2009.* Springer, 2011.

A. M. Etheridge and T. G. Kurtz. Personal communication. *In Preparation*, 2012.

A. M. Etheridge and A. Véber. Personal communication. 2010.

A. M. Etheridge and A. Véber. The spatial Λ-Fleming-Viot process on a large torus:genealogies in the presence of recombination. *Ann. Appl. Probab. (to appear)*, 2012.

S. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.

S. N. Evans. Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincare Probab. Statis.*, B33:339–358, 1998.

W. Ewens. The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, 3:87–112, 1972.

K. Falconer. Random fractals. *Math. Proc. Cambridge Philos. Soc.*, 100(3):559–582, 1986.

K. Falconer. *Fractal Geometry - Mathematical Foundations and Applications (second edition)*. John Wiley, 2003.

D. H. Fearn. Galton-Watson processes with generation dependence. *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.*, 4:159–172, 1972.

L. R. G. Fontes, M. Isopi, C. M. Newman, and K. Ravishankar. The Brownian web: Characterization and convergence. *Ann. Probab.*, 32(4):2857–2833, 2004.

S. Graf. Statistically self-similar fractals. *Probab. Th. Rel. Fields*, 74(3):357–392, 1987.

A. Greven, V. Limic, and A. Winter. Representation theorems for interacting Moran models, interacting Fisher-Wright diffusions and applications. *Electron. J. Probab.*, 10(39):1286, 2005.

G. H. Hardy and M. Riesz. *The general theory of Dirichlet series*. Cambridge University Press, 1915.

T. E. Harris. Coalescing and non-coalescing stochastic flows in $\mathbb{R}$. *Stochastic Process. Appl.*, 17:187–210, 1984.

J. E. Hutchinson. Fractals and self-similarity. *Indiana University Mathematics Journal*, 30 (5):712–747, 1981.

P. Jagers. Galton-Watson processes in varying environments. *J. Appl. Probab.*, 11:174–178, 1974.

M. Jirina. Extinction of non-homogeneous Galton-Watson processes. *J. Appl. Probab.*, 13: 132–137, 1976.

J. Kingman. *Poisson Processes*. Oxford University Press, 1993.

J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.

Y. Le Jan and O. Raimond. Sticky flows on the circle and their noises. *Probab. Th. Rel. Fields*, 129:63–82, 2004.

V. Limic and A. Sturm. The spatial Λ-coalescent. *Electron. J. Probab.*, 11:363–393, 2006.

R. Lyons. Strong laws of large numbers for weakly correlated random variables. *Michigan Math. J.*, 35:353–359, 1988.

R. Lyons. Random walks, capacity and percolation on trees. *Ann. Probab.*, 20(4):2043–2088, 1992.

R. D. Mauldin and S. C. Williams. Random recursive constructions: Asymptotic geometric and topological properties. *Trans. Amer. Math. Soc.*, 295(1):325–346, 1986.

M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29(4):1547–1562, 2001.

P. Mörters. *New perspectives in stochastic geometry*, chapter 8. Oxford University Press, 2010.

J. Neveu. A continuous-state branching process in relation with the GREM model of spin glass theory. *Rapport interne no 267, Ecole Polytechnique*, 1992.

R. Pemantle. Sharpness of second moment criteria for branching and tree-indexed processes. *Classical and modern branching processes, IMA Vol. Math. Appl.*, 84:257–262, 1997.

J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

W. Rudin. *Functional Analysis*. McGraw-Hill, 1991.

H. Saadi. *Λ-Fleming-Viot processes and their spatial extensions*. PhD thesis, Oxford University, 2011.

S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36:1116–1125, 1999.

J. Schweinsberg. A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.*, 5:1–11, 2000a.

J. Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5: 1–50, 2000b.

A. Véber and A. Wakolbinger. New constructions of the spatial Λ-Fleming-Viot process. *In Preparation*, 2012.

A. Véber and F. Yu. Personal communication. 2012.