# Asymptotic Behaviour of Some Measure-Valued Diffusions




*Alison Etheridge*

*New College*

D.Phil. Thesis submitted Trinity term 1989.

# Asymptotic Behaviour of some Measure-valued Diffusions

Alison Mary Etheridge
New College

D.Phil Thesis submitted Trinity Term 1989

## ABSTRACT

Branching diffusions are introduced as a simple model of the growth of a population of rare mutant genes or of the spread of a small epidemic. Measure-valued critical branching diffusion processes are then constructed as the weak limits of the discrete particle branching diffusions when the scales of mass and lifetime are changed in an appropriate way. The distribution of such a measure-valued process is characterized as the unique solution to a martingale problem and this leads to an expression for its Laplace transform in terms of a deterministic dual process.

The dual process is given as the solution to a non-linear second order partial differential equation—the evolution equation of the measure-valued process. It is immediate that those measure-valued processes whose total mass is initially finite suffer extinction. The solution to the evolution equation is written in the form of a series whose coefficients are expressed in terms of the linear semigroup associated with the spatial diffusion part of the process. Estimates on these coefficients lead to a simple proof that if the spatial diffusion part of the measure-valued process is a recurrent (resp. transient) diffusion on $\mathbb{R}^d$ and the initial value of the process is a non-trivial invariant measure of this spatial diffusion then the process has no (resp. has a unique) non-trivial limiting distribution.

In order to overcome the tendency to extinction, a Poisson process of mass creation is introduced into the discrete particle branching diffusion. The resulting processes are rescaled in a way compatible with the previous rescaling to produce a measure-valued process with immigration. This process is also characterized via a deterministic dual, obtained by integrating the solution of the usual evolution equation with respect to time. The series solution to the evolution equation enables the moments to be calculated. It is shown that if the spatial diffusion is positive recurrent then the effect on the limit law of $X(t)/t$ of immigrating at a constant rate is the same as that of conditioning the process without immigration on non-extinction. However, for a null recurrent diffusion it is shown that in the first case the limit is non-trivial (non-random) whereas in the second the limit is zero on any compact region.

Acknowledgement

I should like to thank my supervisor, David Edwards, for directing me towards probability theory and encouraging my efforts in the field. My gratitude is also due to all at McGill University for their hospitality during the academic year 1986/7, especially to John Taylor and Peter March who not only suggested that I should pursue this line of research, but also provided constant support and encouragement. My thanks also go to Chris Rogers for numerous conversations since my return to Britain.

For financial support I am grateful to the Science and Engineering Research Council, the Canadian Rhodes Scholars' Foundation, McGill University and New College, Oxford.

Finally, I am indebted to the Warden, Fellows and students of New College, simply for the pleasure of their company.

Table of Contents

Chapter 1. Introduction ..... 3

Chapter 2. Branching Processes in Biology ..... 8
§2.1. A Problem in Population Genetics ..... 8
§2.2. A Problem in Epidemiology ..... 11

Chapter 3. Finite Particle Branching Diffusions ..... 13
§3.1. Definition of the Process ..... 13
§3.2. The State Space of the Process ..... 16
§3.3. The Generator of the Process ..... 17
§3.4. A Martingale Characterization ..... 21
§3.5. A Remark on the Domain of the Generator ..... 23

Chapter 4. The Measure-valued Processes ..... 25
§4.1. Rescaling the Branching Diffusions ..... 25
§4.2. Relative Compactness of the Rescaled Processes ..... 27
§4.3. The Weak Limit Points : A Martingale Characterization ..... 37
§4.4. Infinite Initial Measures ..... 41
§4.5. Homeomorphisms of $\mathbb{R}^d$ ..... 46

Chapter 5. A Solution to the Evolution Equation ..... 48
§5.1. A Wild Sum ..... 49
§5.2. Asymptotic Estimates : Positive Recurrent Spatial Diffusion ..... 53
§5.3. Asymptotic Estimates : Null Recurrent Spatial Diffusion ..... 59

Chapter 6. Asymptotic Behaviour of the Measure-valued Diffusions ..... 65
§6.1. Processes with Finite Initial Measures ..... 66
§6.2. Infinite Initial Measures : Transient Spatial Motion ..... 67
§6.3. Infinite Initial Measures : Recurrent Spatial Motion ..... 70

§6.4. Regions with Boundaries ..... 74

Chapter 7. Measure-valued Diffusions with Immigration ..... 76

§7.1. Scaling Functions ..... 77

§7.2. Processes with Immigration : Definition and Characterization ..... 78

§7.3. The Total Mass of the Process with Immigration ..... 82

§7.4. Conditioning on Non-extinction : Finite Regions ..... 83

§7.5. Conditioning on Non-extinction : Infinite Regions ..... 86

Chapter 8. Related Results and Open Problems ..... 90

Appendix A. The Weak Topology on $\mathcal{M}_0(\mathbb{R}^d)$ ..... 97

Appendix B. An Existence Theorem ..... 99

Appendix C. The Method of Duality ..... 101

Appendix D. Ordered Binary Rooted Trees ..... 103

Appendix E. The Carleman Condition ..... 105

References ..... 106

## Chapter 1. Introduction

The aim of this thesis is to present some results concerning the asymptotic behaviour of a class of measure-valued branching processes. These processes can be thought of as the extension of the Feller diffusion approximation to the space-time context. We motivate their study by describing two biological models and then present an elementary construction of the measure-valued approximation. This leads to a characterization of the measure-valued process in terms of a martingale problem, and then by applying the 'method of duality' we produce a characterization of its distribution in terms of a second order non-linear partial differential equation—the evolution equation. It is this equation which is our primary tool in the remainder of the thesis. Our study of it reveals that if the initial mass of the measure-valued process is finite, and even in some cases if it is infinite then the process dies out in the limit as time tends to infinity. In trying to overcome this we appeal again to the biological models which suggest that we should be considering a process which has a continuous process of mass creation (which we take to be Poisson). It turns out that the diffusion approximation to a process with this sort of 'immigration' can also be characterized via a deterministic dual process, and that this dual relates in a very natural way to that of the process without immigration—it is obtained by integrating the solution to the evolution equation (multiplied by the rate of immigration) with respect to time. This enables us to exploit further the work for the process without immigration to establish the asymptotic behaviour of the processes with immigration.

In Chapter 2 we describe two of the biological processes which lie behind the study of branching diffusions. A familiarity with the terminology for these processes

makes the reading of the literature a far less formidable task! The origin of the processes can be traced back to the Galton-Watson process (first formulated by the Rev. Henry Watson in (1874)) which was itself designed to model the growth (or rather decline) of English family names. Since then, 'though only after a substantial period of dormancy, this simple branching process has been generalized considerably to better model the spread of a population (nowadays usually of genes rather than gentry!). The individuals in the modern population are no longer confined to reproduce at fixed time intervals, but now each has a random lifetime during which it is allowed to move around some spatial domain. The difficulty with this is that it makes calculations of, for example, transition probabilities extremely complicated. In the case of a large population, Feller (1951) was able to show that by rescaling the parameters of lifetime and mass we can construct a continuous approximation to the size of such a population. This approximation is a diffusion process whose generator can be written down in terms of the mean and variance of the branching mechanisms of the approximating discrete particle processes, and so we have many techniques at our disposal for estimating its value at a given time. In some special cases, Watanabe (1968) extended this notion of 'diffusion approximation' to the space-time context. The processes which this yields are finite measure-valued diffusions. This has gradually been extended by many authors to yield diffusion approximations with infinite initial measures and corresponding to a wide variety of branching diffusions. In Chapter 3 we describe a class of discrete particle branching diffusions to which we are able to find such diffusion approximations. The construction of the measure-valued approximations is given in Chapter 4. In an attempt to make the construction accessible to as wide an audience as possible, we adopt a very direct approach, and, in particular, avoid invoking the heavy machinery of general martingale theory. The

measure-valued diffusions which we construct are characterized via martingale problems. The form of their infinitesimal generators makes these processes especially suited to the application of the 'method of duality'. The proof of the version of duality which we use is straightforward and is given in Appendix C. The dual process which we obtain is the (deterministic) solution to the evolution equation.

The evolution equation is a second order non-linear partial differential equation. It turns out to be easier to consider the equivalent integral equation. Iscoe (1986b) was able to show (using Picard's Theorem) that this integral equation has a unique solution, and if the initial value of the equation lies in the domain of the infinitesimal generator of the spatial diffusion part of the process then the solution to the integral equation also solves the original evolution equation. For the case of a transient spatial diffusion, and for sufficiently small initial values, Dawson (1977) expressed the solution as an infinite sum, the sum being taken over ordered binary rooted trees. Series of this form are well known to physicists who often call them Wild sums (c.f. Chapter 5). In Chapter 5 we present a derivation of this sort of series solution. If the spatial diffusion part of the process is recurrent, then it is not possible to find initial values for which the series solution will be valid for all time. Nonetheless, in Chapter 6, by exploiting the sub-linearity of the evolution equation, we are able to use the series to deduce the asymptotic behaviour of the processes even with recurrent spatial diffusion part. With patience we could actually use the series solution to obtain moments of the measure-valued diffusion process.

Chapter 6 is devoted to establishing the asymptotic behaviour of the processes. A simple calculation tells us that if we start a measure-valued process of this type from a finite initial value then the process will die out. In the special case where the spatial diffusion part of the process is Brownian motion, and the branching is spatially

uniform, Iscoe (1986a) established results about the range of the support. He proved that the process spends its entire 'life' in a ball of finite (random) radius. This result can obviously be used to deduce weaker results about other measure-valued branching processes, but the proof rests heavily on the special symmetry properties of this process. It is this symmetry which makes this particular process (which we call the Dawson process—see Chapter 4) especially amenable to study, for the evolution equation can be reduced to one spatial dimension. Most of Chapter 6 is concerned with the asymptotic behaviour of the processes when they are started from an invariant measure of their spatial diffusion part. This initial value seems to be an obvious place to start looking for stable limiting distributions (if they exist). Here we observe a dichotomy between those processes with transient spatial diffusions and those whose spatial diffusion part is recurrent. We prove that if the spatial diffusion is transient then there is a non-trivial limiting distribution for the measure-valued process. If, however, the spatial diffusion is recurrent, then the process suffers extinction. This result was proved in certain special cases by Dawson (1977), but his proof in the recurrent case required very precise knowledge of the transition density of the spatial diffusion part of the process. The result can be regarded as the continuous analogue of the corresponding result for critical cluster fields which was proved by Kallenberg (1975). Although the result is well known to the experts, its proof for discrete particle branching diffusions seems well buried in the literature! This is perhaps the first illustration that the measure-valued diffusion approximations could be of value to the study of discrete processes—the proof of this result for the measure-valued processes is quite straightforward.

In Chapter 7 we turn to the problem of 'keeping the processes alive'. This is achieved by introducing a Poisson process of immigration. From the biological

viewpoint, this is much more appealing than the more obvious mathematical approach of conditioning the processes on non-extinction. Obviously, we are really interested in the case of recurrent spatial diffusion, and here we observe another division of behaviour, again dictated by the type of spatial diffusion. If we take the distribution of the immigrating mass to be an invariant measure of the spatial diffusion, and immigrate at a constant rate, then we are able to show that if the spatial diffusion is positive recurrent then immigration is, in a sense, equivalent to conditioning on non-extinction. If, however, the spatial diffusion is null recurrent, then it serves to 'smear the mass out' sufficiently that even when the process is conditioned on non-extinction, the amount of mass in any compact set will tend to zero (in distribution). This smearing out effect is sufficiently strong in the case of one-dimensional Brownian motion that the process without immigration will vacate a compact set after a finite random time and never recharge it (with probability one). This result is proved in Iscoe (1986a). In Chapter 8 we will briefly return to the question of stronger convergence properties of the measure-valued processes.

We begin Chapter 8 with a discussion of the effect of dropping some of the regularity assumptions which we make in §3.1 and §5.2. In particular, in §5.2, we impose the condition that the branching part of the processes should be spacially homogeneous. This is largely in attempt to prevent the results of the subsequent chapters becoming obscured by a mass of details. As we explain in Chapter 8, this assumption could be omitted without affecting our results. The rest of the chapter is devoted to a brief survey of some related results and open questions.

In an attempt to keep this thesis as self-contained as possible, we collect together in the appendices some of the results from the theories of weak convergence, binary trees and Markov processes which we cite in the text.

# Chapter 2. Branching Processes in Biology

A primary motivation for the study of branching processes comes from mathematical modelling in biology. We shall therefore devote this chapter to a brief discussion of two of the underlying models and show how they lead to the mathematical problems addressed in subsequent chapters.

Our first model arises in population genetics and our second in the study of epidemics. In both cases, the biological processes are extremely complex so that our models are necessarily greatly oversimplified; nonetheless, they are valuable tools and we shall try to point out some of their strengths as well as their weaknesses.

## §2.1 A problem in population genetics

The characteristics of a human body are determined by its 46 chromosones. These occur in 23 pairs, one of each pair being obtained from each parent. To use an anology of Dawkins (1976), the chromosones are volumes in the architects plans for building the body. They are paired in the sense that to each volume from the father corresponds a volume from the mother and corresponding pages in the two volumes will describe (possibly rival) plans for the same characteristic. We may think of the body as having to form a single hybrid volume from each pair. So in each slot in this new volume a page is copied from one of the two parental volumes. We think of pages as representing genes. Two genes in competition in this way for a slot on a chromosone are said to be alleles of each other. Occasionally a mistake occurs in this process in that, for example, a page or batch of pages is miscopied or misfiled in the new volume. This is then said to be a mutation and a miscopied or misfiled page a

mutant allele.

Suppose that a single mutant gene has appeared in the population. Let c be the expected number of descendants by which it will be represented in the next generation. Then s=logc is defined to be the selective advantage of the mutant. Selection is called disadvantageous, neutral or advantageous according as $s<0$, $s=0$ or $s>0$.

Evolution can be reduced to the process by which some of the genes become more numerous and others less numerous. Two of the principal explanations for the occurrence of this effect are provided by the Darwinist and neutralist theories. A Darwinist attributes evolution to the spread of selectively advantageous mutant types, whereas a neutralist holds a large number of selectively neutral mutants responsible—although the descendants of a single mutant will tend to die out there are sufficiently many mutants around at any one time to explain the tremendous diversity observed in nature.

Anthropologists measure a great number of traits which distinguish human populations and it is observed that individuals of the same 'type' tend to be close to one another. A Darwinist would put this down to geographically varying natural selection—a selective advantage for one gene at one location and for another at another location. The resulting configuration of types is called a cline. A neutralist on the other hand would argue that individuals located close to one another are simply more likely to be related. It seems most likely that in fact both the Darwinist and neutralist processes of evolution contribute to the overall effect that we observe in nature. The problem then is to distinguish the random variation of type of the neutralist from the variation caused by the selective pressure of the Darwinist. We

therefore model 'neutralist evolution'.

Suppose that we have a large 'normal' population in some domain and we introduce a mutation at some point of the domain. We wish to model the growth of the mutant population. We suppose the mutant to move around the domain according to some diffusion process (e.g. Brownian motion). After a random amount of time (assumed to be exponentially distributed) it dies and is replaced at the same location by a random number of offspring (with expectation one since we are modelling selectively neutral types) which go on to diffuse and reproduce in the same way. The presence of the large normal background population makes the probability that two mutants mate with each other negligible so that all mutants may be assumed to behave independently. This then describes a critical branching diffusion process (see Chapter 3).

There is no need to restrict ourselves to introducing a single mutation initially. Indeed, if the background population is sufficiently large, we can, for example, introduce mutants according to a Poisson random field i.e. the expected number of mutants initially in a set $A$ is $Cm(A)$ for some constant $C>0$ where $m$ denotes Lebesgue measure. An interesting special case of this is when the diffusion is Brownian motion and each mutant has 0 or 2 offspring, each with probability $\frac{1}{2}$. Sawyer&Fleischmann (1978) show that in this case the mutant population forms 'clumps' which take over large areas of the domain at high density with regions between clumps becoming empty. The difficulty of this, of course, is that if the density of the mutants in any one region becomes too high, then the primary assumption of a large normal background is no longer satisfied and our model breaks down.

So far, we have only considered a model with a single injection of a number of mutants into the population. In nature one would expect to observe a constant rate of mutations—this leads to what we describe as a process with immigration (see Chapter 7). This type of process can also, as the name suggests, be used to model a population in which mutants move in from outside the domain at a specified rate, for example across the boundaries.

For small populations we must consider these discrete particle models, but for large populations we may consider a continuous approximation achieved by the rescaling which we introduce in Chapter 4. In this way we obtain our measure-valued processes.

## §2.2 A problem in epidemiology

It turns out that mathematically equivalent to the problem discussed above is that of modelling the spread of an epidemic in a large susceptible population subject to the assumptions outlined below (c.f. Frauenthal (1980) Chapter 5).

We imagine that one or several infectious individuals arise in an otherwise disease free area and start the outbreak. Once an individual has contracted the disease, they will remain infectious for an exponentially distributed amount of time. The infectives move around according to a diffusion process and infect other individuals by close contact. It is assumed that there is no latent period for the disease (i.e. that it is transmitted directly on contact), all susceptibles are equally susceptible and all infected individuals are equally infectious. The number of individuals infected by a single infective is governed by a random variable with expectation one. The rôle of the large background of susceptibles is to allow us to ignore the effects of the reduction of the number of individuals at risk of being infected every time a new case

arises.

One of the major advantages of the branching diffusion model is that we can, for example, allow the recovery rate of infected individuals to vary geographically. In the general branching diffusion model we can also allow the expected number of individuals infected by a single infective to vary across the domain. Unfortunately however, in order to use our continuous approximation we must assume that this is everywhere equal to one.

As for our population model, we may wish to introduce a process of 'immigration' of infected individuals into the population. Once public health authorities realize that an epidemic has begun one would expect preventitive measures to be taken so that immigration here may happen at a time-variable (presumably decreasing) rate.

## Chapter 3. Finite Particle Branching Diffusions

We begin this chapter with a precise mathematical description of the finite particle branching diffusion processes which were introduced in Chapter 2. We then characterize the distributions of these processes as solutions to martingale problems. Such a characterization is especially well suited to proving weak convergence theorems (c.f. Ethier&Kurtz (1986) Chapter 4) and will pave the way to an elementary construction of our approximating measure-valued processes in Chapter 4.

Although all our results will be stated for branching diffusions on the whole of $\mathbb{R}^d$, it is clear that they apply equally well on regions $G \subseteq \mathbb{R}^d$ provided that appropriate conditions are imposed on the behaviour of particles on hitting the boundary $\partial G$ of $G$ (e.g. absorption or reflection). However, as we show by example in §3.5, care is needed in describing the domain of the generator of such processes.

### §3.1 Definition of the process

Let $\mathcal{N}$ denote the non-negative integer-valued measures on $\mathbb{R}^d$. A Markov particle system is simply an $\mathcal{N}$-valued Markov process. It is said to be *multiplicative* if its transition probabilities, $\{P_t\}_{t\geq 0}$, satisfy the so-called *branching property*:

$$P_t(\cdot,\nu_1+\nu_2)=P_t(\cdot,\nu_1)*P_t(\cdot,\nu_2) \qquad \forall \nu_1, \nu_2 \in \mathcal{N}, \ \forall t\geq 0$$

(where $*$ denotes convolution). In other words the distribution of the process with initial value $\nu_1+\nu_2$ is equal to that of the sum of two independent copies of the process with initial values $\nu_1$, $\nu_2$ respectively.

Definition (3.1) A *branching diffusion process*, $(Z_t)_{t\geq 0}$, is a multiplicative Markov particle system in which the particles independently undergo random motion and branching described as follows.

(a) Spatial motion Until the time that it branches, each particle (independently of all other particles) performs a diffusion with infinitesimal generator $L$, say, in $\mathbb{R}^d$.

(b) The branching rate The probability that a particle located at $x\varepsilon\mathbb{R}^d$ at time $t$ branches in the interval $[t,t+\delta t)$ is $V(x)\delta t+o(\delta t)$. $V(x)$ is called the *branching rate.*

(c) The branching mechanism When a particle 'branches', it dies and is replaced at the same location by a random number of offspring governed by what we call the branching mechanism. We denote the generator of the branching mechanism by $\Phi(s,x)$; i.e. if $p_k(x)$ is the probability that a particle which dies at $x$ gives rise to $k$ offspring, then

$$\Phi(s,x)=\sum_{k=0}^{\infty} p_k(x)s^k \qquad\qquad x\varepsilon\mathbb{R}^d,\ s\varepsilon\mathbb{R}$$

Notice that, by altering $V(x)$ if necessary, we may as well assume that $p_1(x)\equiv 0$.

We shall call this the $(L,V\Phi)-$*branching diffusion (process)* and refer to the diffusion generated by $L$ as the *spatial diffusion part* or *base diffusion.*

Recall that a function $l(t)$ on $(0,\infty)$ is said to be slowly varying at infinity if for every fixed $x$,

$$\frac{l(tx)}{l(t)}\to 1 \qquad\qquad \text{as } t\to\infty$$

For the remainder of this thesis we shall make the following assumptions:

(i) $V(x) \varepsilon C_b^+(\mathbb{R}^d)$

(ii) $p_k(x)$ is a continuous function of $x$ for each $k$.

(iii) $$0 < \inf_x \frac{\partial^2}{\partial s^2}\Phi(s,x)\Big|_{s=1} \leq \sup_x \frac{\partial^2}{\partial s^2}\Phi(s,x)\Big|_{s=1} < \infty$$

(iv) $L$ is a uniformly strictly elliptic second order differential operator of the form

$$Lf(x) = \sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i} \qquad f \varepsilon C^2(\mathbb{R}^d)$$

where $a_{ij}$ are bounded continuous and $b_i$ are bounded measurable functions on $\mathbb{R}^d$.

(v) If the diffusion generated by $L$ is recurrent, then $\exists$ a function $h(t)$ on $(0,\infty)$, and an invariant measure $\mu$ of the diffusion with

$$\frac{1}{h(t)}\int_0^t T_s \phi \, ds \to \int \phi \, d\mu \qquad \text{as } t \to \infty$$

and

$$h(t) \sim t^{\rho} l(t) \qquad \text{as } t \to \infty$$

for some slowly varying function $l(t)$ and some $0 \leq \rho \leq 1$.

Remark Condition (v) is not used anywhere in the construction of the measure-valued processes, but it is essential for our method of establishing asymptotic behaviour. In fact it will always hold when the diffusion is positive recurrent. We defer the discussion of how our results extend to cases where some of (i)−(iv) are dropped to Chapter 8.

An example which we repeatedly meet in the literature is the following:

Definition 3.2 The *simple branching diffusion process on* $\mathbb{R}^d$ is that process for which $L=\frac{1}{2}\Delta$ ($\Delta$ the Laplacian on $\mathbb{R}^d$), $V(x)\equiv C$ for some strictly positive constant $C$ and $\Phi(s,x)\equiv\frac{1}{2}(1+s^2)$.

For obvious reasons this process is also sometimes referred to as *critical branching Brownian motion.*

§3.2 The state space of the process

By definition, $Z(t)$ is an $\mathcal{N}$-valued Markov process. For the purpose of defining a topology on $\mathcal{N}$ and calculating the generator of $Z(t)$, it is convenient to think of $\mathcal{N}$ as follows.

Let $(\mathbb{R}^d)^{(n)}$ be defined for $n\geq 1$ to be the space of equivalence classes of n-tuples of elements $x\varepsilon\mathbb{R}^d$ modulo permutations. Let $\pi^{(n)}\colon (\mathbb{R}^d)^n\to(\mathbb{R}^d)^{(n)}$ be the map taking $(x_1,\ldots,x_n)$ in $(\mathbb{R}^d)^n$ into its equivalence class $\{x_1,\ldots,x_n\}$ in $(\mathbb{R}^d)^{(n)}$ and topologize $(\mathbb{R}^d)^{(n)}$ so that $\pi^{(n)}$ is both open and continuous.

Suppose now that $\{\emptyset\}$ is an abstract point. Set

$$E_1=\{\emptyset\}\cup\bigcup_{n=1}^{\infty}(\mathbb{R}^d)^{(n)}$$

and topologize $E_1$ in such a way that $\{\emptyset\}$ and $(\mathbb{R}^d)^{(n)}$ for each $n\geq 1$ are both open and closed. Then we may identify $\mathcal{N}$ with $E_1$ and this defines a topology on $\mathcal{N}$.

It will be important in Chapter 4 that the topology defined above can in fact be constructed as follows. Let $\mathcal{M}_0(\mathbb{R}^d)$ be the space of finite positive Borel measures on $\mathbb{R}^d$ with the weak topology i.e.

$$\lim_{n\to\infty}\mu_n=\mu \quad iff \quad \lim_{n\to\infty}\int f d\mu_n=\int f d\mu \qquad \forall f\in C_b(\mathbb{R}^d)$$

The weak topology is metrizable by a natural extension, $\rho$, of the Prohorov metric on probability measures. (For the definition of the Prohorov metric and this result, see Appendix A.) $\rho$ is just the completion of the following metric on non-zero elements of $\mathcal{M}_0(\mathbb{R}^d)$

$$\rho(\mu,\nu)=\tilde{\rho}\left(\frac{\mu}{|\mu|},\frac{\nu}{|\nu|}\right)+\big||\mu|-|\nu|\big| \qquad \mu\neq 0\neq\nu$$

where $\tilde{\rho}$ is the Prohorov metric and $|\mu|=\int 1 d\mu$. We write $\overline{\mathcal{M}_0(\mathbb{R}^d)}^{\rho}$ for the metric space obtained by completing $\mathcal{M}_0(\mathbb{R}^d)$ in this way. We will always identify the elements $\mu$ of the completion for which $|\mu|=0$ in what follows. This is, of course, compatible with the weak topology.

### §3.3 The generator of the process

We now derive Skorohod's equation for the $(L,V\Phi)$-branching diffusion and use it to calculate the generator of the process. It is assumed that $(Z(t))_{t\geq 0}$ is a right

continuous strong Markov process.

If $f \varepsilon C(\mathbb{R}^d)$ with $0<f<1$, then we extend $f$ to $\mathcal{N}$ by

$$f\{\emptyset\}=1$$
$$f(\{x_1,\ldots,x_n\})=\prod_{i=1}^{n} f(x_i) \tag{3.1}$$

Notice that the functions $f$ on $\mathcal{N}$ form a determining set. The distribution of a Markov process is completely determined by its one-dimensional conditional distributions and so by virtue of the branching property, to characterize a branching diffusion process it is sufficient to compute

$$\hat{T}_t f(x)=\mathbb{E}[f(Z_t)|Z_0=\{x\}\varepsilon(\mathbb{R}^d)^{(1)}]$$

Let $\tau$ denote the random time at which a single particle initially in the system first branches. We write $\mathcal{F}_\tau=\sigma\{Z_u\colon 0\leq u<\tau\}$ (the minimal $\sigma$-field with respect to which $Z_u$ is measurable for all $u<\tau$. Also let $\overline{X}_t$ be the base diffusion of the process killed at time $\tau$ and write

$$\overline{T}_t f(x)=\mathbb{E}[f(\overline{X}_t)|\overline{X}_0=x]=\mathbb{E}_x f(\overline{X}_t)$$

for the corresponding semigroup.

Suppose that $\chi_t$ is defined by

$$\chi_t=\begin{cases} 1 & t\leq\tau \\ 0 & t>\tau \end{cases}$$

Then we may write

$$f(Z_t)=f(\overline{X}_t)\chi_t+f(Z_t)(1-\chi_t)$$

Clearly,

$$\mathbb{E}[f(\overline{X}_t)\chi_t|Z_0=\{x\}\varepsilon(\mathbb{R}^d)^{(1)}]=\overline{T}_t f(x)$$

If $Z_\tau\varepsilon(\mathbb{R}^d)^{(k)}$, then we denote the process after time $\tau$ by $(Z_t^{(1)},\ldots, Z_t^{(k)})$ where $\{Z_t^{(i)}: t>\tau\}_{i=1}^k$ are independent and $Z_\tau^{(i)}=(\overline{X}_\tau)\ \forall i$. With this notation

$$\mathbb{E}[f(Z_t)(1-\chi_t)|Z_0=\{x\}\varepsilon(\mathbb{R}^d)^{(1)}]=\mathbb{E}\Big(\mathbb{E}[f(Z_t)|\mathcal{F}_\tau](1-\chi_t)|Z_0=\{x\}\varepsilon(\mathbb{R}^d)^{(1)}\Big)$$

$$=\mathbb{E}\left(\sum_{k=0}^{\infty}p_k(\overline{X}_\tau)\cdot\prod_{i=1}^{k}\mathbb{E}[f(Z_t^{(i)})|Z_\tau^{(i)}=\overline{X}_\tau](1-\chi_t)|Z_0=\{x\}\varepsilon(\mathbb{R}^d)^{(1)}\right)$$

$$=\mathbb{E}\left(\Phi\Big(\hat{T}_{t-\tau}f(\overline{X}_\tau),\ \overline{X}_\tau\Big)(1-\chi_t)|Z_0=\{x\}\varepsilon(\mathbb{R}^d)^{(1)}\right)$$

(by the independence of $\{Z_t^{(i)}: t>\tau\}_{i=1}^k$ and the strong Markov property).

Now let $K_x(du,dy)$ be the measure on $[0,\infty)\times\mathbb{R}^d$ given by

$$\mathbb{E}_x[g(\tau,\overline{X}_\tau)]=\int_{[0,\infty)}\int g(u,y)K_x(du,dy)\qquad\qquad\forall\, g\varepsilon\, C_b([0,\infty)\times\mathbb{R}^d)$$

The expression on the right hand side above then becomes

$$\int_0^t\int\Phi\Big(\hat{T}_{t-u}f(y),y\Big)K_x(du,dy)$$

(The $(1-\chi_t)$ term causes the integrand to vanish for $u>t$.) We then obtain the following integral equation for $\hat{T}_t f(x)$

$$\hat{T}_t f(x) = \overline{T}_t f(x) + \int_0^t \int \Phi\Big(\hat{T}_{t-u} f(y), y\Big) K_x(du, dy)$$

This is Skorohod's equation (first obtained by Skorohod (1964)).

By means of successive approximation it can be shown that this equation has a unique solution for sufficiently small t, giving the transition probabilities of the process for such t, and hence by the Markov property for all t.

It is now straightforward to calculate the infinitesimal generator, $\mathcal{L}_1$, of the $(L, V\Phi)$-branching diffusion on the extension to $\mathcal{N}$ of functions $f \varepsilon C_b(\mathbb{R}^d) \cap \mathfrak{D}(L)$ with $0<f<1$

$$\mathcal{L}_1 f(\{x\}) = \lim_{t \downarrow 0} \left( \frac{\hat{T}_t f(x) - f(x)}{t} \right)$$

$$= \overline{L} f(x) + V(x) \Phi\Big(f(x), x\Big)$$

$$= \Big(L - V(x) I\Big) f(x) + V(x) \Phi\Big(f(x), x\Big) \quad \text{(Dynkin (1965) p.11)}$$

$$= L f(x) + V(x) \Big(\Phi\Big(f(x), x\Big) - f(x)\Big)$$

and so using the branching property,

$$\mathcal{L}_1 f(\{x_1 \ldots, x_n\}) = \sum_{j=1}^{n} \Big\{ L f(x_j) + V(x_j) \Big(\Phi\Big(f(x_j), x_j\Big) - f(x_j)\Big) \Big\} \prod_{\substack{i=1 \\ i \neq j}}^{n} f(x_i)$$

Now notice that

$$\prod_{i=1}^{n} f(x_i) = exp(<logf, \sum_{i=1}^{n} \delta_{x_i}>)$$

Thus we can also write the generator of our process as follows

$$\mathcal{L}_1(exp<logf,\mu>) = \left\langle \frac{Lf + V(\Phi(f)-f)}{f}, \mu \right\rangle exp(<logf,\mu>)$$

where $\Big(\Phi(f)\Big)(x) = \Phi\Big(f(x),x\Big)$.

Remark We have tacitly assumed that our branching processes do actually exist. However, it is not difficult to see that under our conditions (i)−(iv) the semigroup $\hat{T}_t$ corresponds to a normal transition function satisfying the conditions given in Appendix B to ensure that $\hat{T}_t$ really is the transition semigroup of a strong Markov process. For a construction of more general branching Markov processes the reader is referred to Ikeda, Nagasawa&Watanabe (1968) part II.

## §3.4 A martingale characterization

As we mentioned before, for our purposes it is best to characterize the $(L, V\Phi)-$branching diffusion in terms of a martingale problem.

For a metric space $(E,\rho)$ we write $D_E([0,\infty))$, or sometimes $D([0,\infty),E)$, for the space of càdlàg $E$-valued functions on $[0,\infty)$ (i.e. right continuous $E$-valued functions with left limits). Let $\Omega_1$ denote $D_{\mathcal{N}}([0,\infty))$. Since $\mathcal{N}$ is a metric space we may endow

$\Omega_1$ with the usual Skorohod topology (see e.g. Ethier&Kurtz (1986) Chapter 5 §3).

We write $\eta(\cdot)$ for a generic element of $\Omega_1$. For $t\geq 0$, set $\mathcal{F}_t=\sigma\{\eta(s) : 0\leq s\leq t\}$ and $\mathcal{F}=\sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$. Suppose that, for $\mu\in\mathcal{N}$, $\mathbb{P}^\mu$ denotes the distribution of the $(L,V\Phi)$-branching diffusion started from $\mu$ and $\mathbb{E}^\mu$ the corresponding expectation operator. Then for $F\in\mathcal{D}(\mathcal{L}_1)$

$$\frac{d}{ds}\mathbb{E}^{\eta(0)}[F(\eta(s))]=\mathbb{E}^{\eta(0)}[\mathcal{L}_1F(\eta(s))]$$

so that

$$\mathbb{E}^{\eta(0)}[F(\eta(t+u))-F(\eta(t))]=\mathbb{E}^{\eta(0)}\Big(\int_t^{t+u}\mathcal{L}_1F(\eta(s))ds\Big)$$

and then using the Markov property

$$\mathbb{E}^{\eta(0)}[F(\eta(t+u))-F(\eta(t))|\mathcal{F}_t]=\mathbb{E}^{\eta(t)}[F(\eta(u))-F(\eta(0))]$$

$$=\mathbb{E}^{\eta(t)}\Big(\int_0^{u}\mathcal{L}_1F(\eta(s))ds\Big)$$

$$=\mathbb{E}^{\eta(0)}\Big(\int_t^{t+u}\mathcal{L}_1F(\eta(s))ds\Big|\mathcal{F}_t\Big)$$

i.e. $F(\eta(t))-\int_0^t\mathcal{L}_1F(\eta(s))ds$ is a $\mathbb{P}^{\eta(0)}$−martingale for all $F\in\mathcal{D}(\mathcal{L}_1)$.

**Definition 3.3** Given $\nu\in\mathcal{N}$, we say that the probability measure $\mathbb{P}$ on $(\Omega_1,\mathcal{F})$ solves the martingale problem for $\mathcal{L}_1$ started from $\nu$ (or more briefly for $(\mathcal{L}_1,\nu)$) if

$$\mathbb{P}[\eta(0)=\nu]=1$$

and $F(\eta(t))-\int_0^t \mathcal{L}_1 F(\eta(s))ds$ is a $\mathbb{P}$-martingale for all $F\varepsilon\mathcal{D}(\mathcal{L}_1)$.

We have seen that the distribution of the $(L,V\Phi)-$branching diffusion starting from $\nu$ solves the $(\mathcal{L}_1,\nu)-$martingale problem. In fact, much more is true.

Theorem 3.4 For each $\nu\varepsilon\mathcal{N}$ there is exactly one solution to the $(\mathcal{L}_1,\nu)-$martingale problem.

Thus this characterizes the distribution of our branching diffusions uniquely. We do not use uniqueness in what follows and so we do not prove it here. However, the proof is entirely parallel to the corresponding proof of uniqueness of the solution to the martingale problem for the measure-valued processes in §4.3.

§3.5 A remark on the domain of the generator

A branching diffusion process (like any other diffusion process) is completely determined by its full generator. We should like to emphasize the importance of the domain of the generator. We illustrate by an example.

Consider the simple branching diffusion process on a domain $G\subseteq\mathbb{R}^d$. The generator $\mathcal{A}$ is of the form

$$\mathcal{A}exp(<logf,\mu>)=\left\langle \frac{\Delta f+(1-f)^2}{2f},\ \mu\right\rangle exp(<logf,\mu>)$$

We take the following two domains for $\mathcal{A}$

$$\mathcal{D}^1(\mathcal{A})=\{exp(<logf,\mu>)\ :\ f\varepsilon\mathcal{D}(L),\ f(x)\to C\ as\ x\to\partial G\ some\ C\varepsilon\mathbb{R}^+\}$$

$$\mathfrak{D}^2(\mathcal{A})=\{exp(<logf,\mu>) : f\varepsilon\mathfrak{D}(L),\ f(x)\to 1\ as\ x\to\partial G\}$$

In both cases the spatial diffusion part is absorbed on $\partial G$. Notice, however, that in the first case a particle continues to branch on the boundary whereas in the second case once on the boundary a particle will cease to branch.

## Chapter 4. The Measure-valued Processes

Feller (1951) was the first to observe that although if we are modelling the growth of a small population (or epidemic) we must use a discrete particle system, for large populations we can apply a continuous approximation. This leads to a diffusion process. The approximation is obtained by rescaling so that parameters such as birth rate, death rate and 'mass' of each particle tend to zero or infinity. We are concerned with the extension of this idea to the space-time context. In this way we obtain a measure-valued limit process.

Such processes have been studied by many authors. In particular, Dawson, Iscoe, Sugitani (and others) have studied in great detail the process obtained as the diffusion approximation to the simple branching diffusion process. Here we shall refer to this process as the *Dawson process* (even though it was first constructed by Watanabe (1968)!).

In this chapter we construct and characterize the diffusion approximations to those $(L, V\Phi)-$branching diffusions with critical branching mechanism (i.e. for which $\frac{\partial \Phi}{\partial s}(s,x)\Big|_{s=1} \equiv 1$).

### §4.1 Rescaling the branching diffusions

Suppose that we wish to construct the Dawson process started from $\delta_x$ (the unit point mass at $x \varepsilon \mathbb{R}^d$). Then we rescale the simple branching diffusion process as follows (Watanabe (1968)) : Let $Z^n(t)$ denote the simple branching diffusion with branching rate $V_n = n$ starting from $n$ particles located at $x$. Then the Dawson process started from $\delta_x$ is the weak limit as $n \to \infty$ of the processes $X^n(t) = \frac{1}{n} Z^n(t)$.

The critical branching mechanism is clearly essential to convergence of these processes to a non-trivial limit (the limit of a such a sequence of rescaled supercritical process would instantly explode, that of a sequence of subcritical process would instantly die). We use the same rescaling to construct our more general diffusion approximations. Thus, let $Z^n(t)$ denote now the $(L,nV\Phi)-$branching diffusion process and write $X^n(t)=\frac{1}{n}Z^n(t)$.

We should like to be able to start our measure-valued process from arbitrary $\nu\varepsilon\mathcal{M}_0(\mathbb{R}^d)$. Initially, however, we restrict our attention to $\nu$ of compact support. We will be able to drop this restriction later (§4.4). Initially then, we choose $\{Z^n(0)\}_{n\geq 1}$ in such a way that $\{\frac{1}{n}Z^n(0)\}_{n\geq 1}$ converges weakly as $n\rightarrow\infty$ to a finite measure $\nu$ of compact support. We may further insist that

$$\text{(4.1)} \qquad \begin{cases} |Z^n(0)|\leq Cn \\ supp Z^n(0)\subseteq K_0 \end{cases} \qquad \forall n$$

for some fixed $C>0$ and some fixed compact set $K_0\subseteq\mathbb{R}^d$. (This will be essential in our proof of Proposition 4.3.)

The state space for $X^n$ is obviously

$$E_n=\{\mu\varepsilon\mathcal{M}_0(\mathbb{R}^d) : \mu=\frac{1}{n}\sum_{i=1}^{k}\delta_{x_i}\ ,\ x_i\varepsilon\mathbb{R}^d,\ k\varepsilon\mathbb{N}\}$$

and from our expression for $\mathcal{L}_1$, we may write down the generator, $\mathcal{L}_n$, of the process $X^n$, acting on functions of the form $exp(<nlogf,\cdot>)$ where $f\varepsilon\mathcal{D}(L)$, $0<f\leq 1$. On these functions we have

$$\mathcal{L}_n(exp(<nlogf,\mu>)=\left\langle \frac{nLf+n^2V(\Phi(f)-f)}{f}\ ,\ \mu\right\rangle exp(<nlogf,\mu>)$$

As in Chapter 3, we can actually characterize $X^n$ with $X^n(0)=\nu_n$ as that process whose distribution is the unique solution to the martingale problem for $(\mathcal{L}_n,\nu_n)$ over $D([0,\infty),E_n)$ (although we will not use uniqueness in what follows).

## §4.2 Relative compactness of the rescaled processes

Suppose then that we have been given a weakly convergent sequence $\{\nu_n=\frac{1}{n}Z^n(0)\}_{n\geq 1}$ satisfying the conditions (4.1). Our aim in this section is to prove that $\{X^n(0)\}_{n\geq 1}$ is then relatively compact in the weak topology. $\overline{\mathcal{M}_0(\mathbb{R}^d)}^{\rho}$ is a complete separable metric space (see Appendix A). We should like therefore to apply the following theorem (due essentially to Prohorov)

Theorem 4.1 (see e.g. Ethier&Kurtz (1986) Chapter 3 §7.)

Let $(E,\rho)$ be a complete separable metric space. For $x(s)\varepsilon D_E([0,\infty))$ and fixed $T>\delta>0$ define

$$w'(x,\delta,T)=\inf_{\{t_i\}}\max_i \sup_{s,t\varepsilon[t_{i-1},t_i)}\rho(x(s),x(t))$$

where $\{t_i\}$ ranges over all partitions of the form

$$0=t_0<t_1<\ldots<t_{n-1}<t_m=T$$

with $m\geq 1$ and

$$\min_{1\le i\le m}(t_i - t_{i-1}) > \delta$$

Let $\{X^n\}_{n\ge 1}$ be a sequence of processes with sample paths in $D_E([0,\infty))$. Then $\{X^n\}_{n\ge 1}$ is relatively compact *iff* the following two conditions hold:

(a) For every $\eta>0$ and every (rational) $t\ge 0$, $\exists$ a compact set $\Gamma^{\eta}_{\eta,t}\subseteq E$ s.t.

$$\liminf_{n\to\infty} \mathbb{P}[X^n(t)\varepsilon\Gamma^{\eta}_{\eta,t}]\ge 1-\eta$$

(b) For every $\eta>0$ and $T>0$, $\exists$ $\delta>0$ s.t.

$$\limsup_{n\to\infty} \mathbb{P}[w'(X^n,\delta,T)\ge\eta]\le\eta$$

(For a compact set $K\subseteq E$, $K^{\epsilon}=\{x\varepsilon E : \inf_{s\epsilon K}\rho(x,y)<\epsilon\}$.)

Condition (a) is the expected requirement of control of size and range of the processes; (b) prevents large 'jumps' from being too close together—it implies that for any $\eta>0$ we can find $N$ so that (with high probability) none of the processes $X^n$, $n\ge N$, has a limit point of jumps of size greater than $\eta$.

Unfortunately, this is where we pay the price for choosing the metric $\rho$. Although, as we shall see below, condition (a) is easily verified, condition (b) may fail when the process $X^n$ becomes too small. The difficulty is that if, for example, $x_1$, $x_2$ are a positive distance, $r$ say, apart in $\mathbb{R}^d$, then for every $n$,

$$\rho(\tfrac{1}{n}\delta_{x_1},\tfrac{1}{n}\delta_{x_2})=r\wedge 1$$

The effect of this is that our metric takes account of how the processes die and distinguishes between 'different zero-masses'. This won't, however, prove to be an obstacle to relative compactness in the weak topology since there, there is only one zero!

Let $\tau_\alpha(t)=sup\{s\leq t : |X^n(s)|\geq\alpha\}$ and write

$$X^n_\alpha(t)=X^n(t\wedge\tau_\alpha)$$

We begin by showing that $\{X^n_\alpha(t)\}_{n\geq 1}$ is relatively compact in $\overline{\mathcal{M}_0(\mathbb{R}^d)}^\rho$ for each fixed $\alpha>0$.

From our expression for the metric $\rho$, it is clear that condition (a) of Theorem 4.1 will certainly be satisfied by $\{X^n_\alpha(t)\}_{n\geq 1}$ if we can prove the following

(a)$'$ Given $t\geq 0$ and $\epsilon>0$

(i) $\exists B=B(t,\epsilon)\in\mathbb{R}_+$ s.t.

$$\mathbb{P}[|X^n_\alpha(t)|\leq B]>1-\epsilon \qquad \forall n$$

(ii) $\exists$ a compact set $K=K(t,\epsilon)\subseteq\mathbb{R}^d$ s.t.

$$\mathbb{P}[\left\langle K^c, \frac{X^n_\alpha(t)}{|X^n_\alpha(t)|}\right\rangle<\epsilon]>1-\epsilon \qquad \forall n$$

where $K^c$ denotes the complement of $K$ in $\mathbb{R}^d$.

Remark Condition (ii) is the condition that

$$\left\{\frac{X_\alpha^n(t)}{|X_\alpha^n(t)|}\right\}_{n\geq 1}$$

is relatively compact in the completion of the set of probability measures on $\mathbb{R}^d$ with respect to the Prohorov metric. See Appendix A.

Indeed, from our definition of $X_\alpha^n$ it obviously suffices to prove (a)$'$ for the processes $\{X^n\}_{n\geq 1}$.

Our assumption on the spatial diffusion part of the process (see p.15) implies that if a particle diffusing according to $L$ starts off in the compact set $K_0$ at time zero, then after any finite time $t$, with high probability the particle lies within another compact set (depending on $t$). Condition (ii) then follows easily. Condition (i) is immediate from the following result which (for easy reference) we state as a lemma.

Lemma 4.2 Let $\mathbb{P}^n$ denote the distribution of $X^n$. Then

$$\mathbb{P}^n[\sup_t|\eta(t)| > x] \leq \frac{1}{x}|\nu_n|$$

Proof Since the branching mechanism of $X^n$ is critical, $|\eta(t)|$ is a $\mathbb{P}^n$-martingale (for all $n$).

To see that $\{X_\alpha^n\}_{n\geq 1}$ is relatively compact in the weak topology, it remains to verify condition (b) of Theorem 4.1.

**Proposition 4.3** Suppose that $T>0$ is fixed. Then for each $\eta>0$, $\exists\delta>0$ and $N\epsilon\mathbb{N}$ such that for each $n\geq N$ there is a partition

$$0=t_0<t_1<\ldots<t_m=T \text{ of } [0,T]$$

with $\min_i (t_i-t_{i-1})>\delta$ for which

$$\max_i \sup_{s,t\epsilon[t_{i-1},t_i)} \mathbb{P}[\rho\Big(X_\alpha^n(s),X_\alpha^n(t)\Big)\geq\eta]\leq\eta$$

**Proof** We will prove this via two lemmas. Recall that

$$\rho\Big(X_\alpha^n(s),X_\alpha^n(t)\Big)=\tilde{\rho}\left(\frac{X_\alpha^n(s)}{|X_\alpha^n(s)|},\frac{X_\alpha^n(t)}{|X_\alpha^n(t)|}\right)+\|X_\alpha^n(s)|-|X_\alpha^n(t)\|$$

It will therefore suffice to prove the corresponding result for the two separate parts of the metric.

**Lemma 4.4** Suppose that $T>0$ is fixed. Then for each $\epsilon>0$, $\exists\delta_1>0$ such that $|t-s|<2\delta_1$ implies

$$\mathbb{P}[||X_\alpha^n(s)|-|X_\alpha^n(t)||>\epsilon]<\epsilon$$

Lemma 4.5 Suppose that $T>0$ is fixed. Then for each $\epsilon>0$, $\exists\delta_2>0$ such that

$|t-s|<2\delta_2$ implies

$$\mathbb{P}[\tilde{\rho}\left(\frac{X_\alpha^n(s)}{|X_\alpha^n(s)|},\frac{X_\alpha^n(t)}{|X_\alpha^n(t)|}\right)\geq\epsilon]\leq\epsilon$$

Having proved these two lemmas we just set $\delta=min(\delta_1,\delta_2)$ and $\epsilon=\eta/2$.

Proof of Lemma 4.4 By our construction of $X_\alpha^n$, it is certainly the case that given $\epsilon>0$,

$$\mathbb{P}[\|X_\alpha^n(s)|-|X_\alpha^n(t)\|>\epsilon]\leq\mathbb{P}[\|X^n(s)|-|X^n(t)\|>\epsilon]$$

Consider then $\|X^n(s)|-|X^n(t)\|$

$$\mathbb{P}[\|X^n(s)|-|X^n(t)\|>\epsilon]=\mathbb{P}[\|X^n(s)|-|X^n(t)\|^2>\epsilon^2]$$

$$\leq\frac{1}{\epsilon^2}\mathbb{E}[\|X^n(s)|-|X^n(t)\|^2]$$

$$=\frac{1}{n^2\epsilon^2}\mathbb{E}[\|Z^n(s)|-|Z^n(t)\|^2]$$

Write $Z_1^n(t,x)$ for the branching diffusion process with branching rate $nV$ and $Z_1^n(0,x)=\delta_x$, $x\varepsilon\mathbb{R}^d$, and $W(k)$ for the $k^{th}$ generation of some Galton-Watson process whose variance $\bar{\sigma}$ is given by

$$\bar{\sigma}+1=\sup_x\sum_{k=1}^{\infty}k^2p_k(x)$$

(i.e. the variance of W is the supremum of the variance of $Z^n$.) Also let $[k,[s,t)]_x$ denote the event that a single particle located at $x$ at time $s$ has $k^{th}$ generation (and no higher than $k^{th}$ generation) offspring in the time interval $[s,t)$. Now the right hand side above is just

$$\frac{1}{n^2\epsilon^2}\mathbb{E}\Big(\mathbb{E}[\|Z^n(s)|-|Z^n(t)\|^2|\ |Z^n(s)|]\Big)$$

$$=\frac{1}{n^2\epsilon^2}\mathbb{E}\Big(\sum_j\mathbb{E}[\|Z^n(s)|-|Z^n(t)\|^2|\ |Z^n(s)|=j]\mathbb{P}[|Z^n(s)|=j]\Big)$$

$$\leq\frac{1}{n^2\epsilon^2}\mathbb{E}\Big(\sum_j j\cdot\sup_x\mathbb{E}[\|Z_1^n(s,x)|-|Z_1^n(t,x)\|^2]\mathbb{P}[|Z^n(s)|=j]\Big)$$

$$=\frac{1}{n^2\epsilon^2}\mathbb{E}[|Z^n(s)|]\cdot\sup_x\mathbb{E}[\|Z_1^n(s,x)|-|Z_1^n(t,x)\|^2]$$

$$=\frac{1}{n^2\epsilon^2}|Z^n(0)|\cdot\sup_x\Big\{\sum_{k=0}^{\infty}\mathbb{E}\Big(\mathbb{E}[\|Z_1^n(s,x)|-|Z_1^n(t,x)\|^2|\ [k,[s,t)]_x]\cdot\mathbb{P}^n[[k,[s,t)]_x]\Big\}$$

$$\leq\frac{1}{n^2\epsilon^2}|Z^n(0)|\sum_{k=0}^{\infty}varW(k)\sup_x\mathbb{P}^n[[k,[s,t)]_x]$$

$$=\frac{1}{n^2\epsilon^2}|Z^n(0)|\sum_{k=0}^{\infty}kvarW(1)\sup_x\mathbb{P}^n[[k,[s,t)]_x]$$

$$=\frac{1}{n^2\epsilon^2}\ |Z^n(0)|\bar{\sigma}\sup_x\Big\{\sum_{k=0}^{\infty}k\mathbb{P}^n[[k,[s,t)]_x]\Big\}$$

$$=\frac{1}{n\epsilon^2}\ |Z^n(0)|\bar{\sigma}\sup_x\Big\{\sum_{k=0}^{\infty}k\mathbb{P}^1[[k,[s,t)]_x]\Big\}$$

Now, by our (temporary) assumptions, $|\mathrm{Z}^n(0)|\leq Cn$ for some $C>0$ and (by the assumptions of §3.1) $\bar{\sigma}$ is bounded. Also, $\overline{V}\equiv\sup_x V(x)<\infty$ and

$$\sup_x\Big\{\sum_{k=0}^{\infty}k\mathbb{P}^1[[k,[s,t)]_x]\Big\}\leq\overline{\mathrm{V}}|t-s|$$

It is clear then that given $\epsilon>0$, we may choose $\delta_1=\delta_1(\epsilon)>0$ so that $|t-s|<2\delta_1\Rightarrow$

$$\mathbb{P}[\|X^n(s)|-|X^n(t)\|>\epsilon]<\epsilon$$

Proof of Lemma 4.5 We consider events of the form

$$A_\epsilon^n=\left[\tilde{\rho}\left(\frac{X_\alpha^n(s)}{|X_\alpha^n(s)|}\ ,\ \frac{X_\alpha^n(t)}{|X_\alpha^n(t)|}\right)>\epsilon\right]$$

By definition of the Prohorov metric, the event $A_\epsilon^n$ is excatly the event that there exists a closed set $F\subseteq\mathbb{R}^d$ such that

$$\left\langle\chi_F,\frac{\mathrm{X}_\alpha^n(s)}{|X_\alpha^n(s)|}\right\rangle>\left\langle\chi_{F^\epsilon},\frac{\mathrm{X}_\alpha^n(t)}{|X_\alpha^n(t)|}\right\rangle+\epsilon$$

Fix one such $F$. Let $\{X_F^n(u)\}_{u\geq s}$ denote that portion of the process $X^n$ (after time $s$) which evolves from those particles of $X^n(s)$ which lie in the set $F$, with the corresponding meaning for $\{X_{\alpha,F}^n(u)\}_{u\geq s}$. Then $\mathbb{P}(A_\epsilon^n)$ is certainly no greater than

$$\mathbb{P}\left[\left\langle \chi_F, \frac{X^n_\alpha(s)}{|X^n_\alpha(s)|}\right\rangle > \left\langle \chi_{F^\epsilon}, \frac{X^n_{\alpha,F}(t)}{|X^n_\alpha(t)|}\right\rangle + \epsilon\right]$$

Now in a similar way to the proof of (a)′ (ii) on p.29, the assumptions on the spatial diffusion part of our processes imply that given $\epsilon>0$ we can choose $\delta_0=\delta_0(\epsilon)>0$ so that $|t-s|<2\delta_0 \Rightarrow$

$$\mathbb{P}\left[\left\langle \chi_{F^\epsilon}, \frac{X^n_F(t)}{|X^n_F(t)|}\right\rangle > 1-\frac{\epsilon}{2}\right] > 1-\frac{\epsilon}{2}$$

The same is then obviously true of $X^n_{\alpha,F}(t)$.

The proof of the lemma will be complete if we can show that given $\epsilon>0$, we can choose $\delta_0'=\delta_0'(\epsilon)>0$ so that $|t-s|<2\delta_0' \Rightarrow$

$$\mathbb{P}\left[\left|\frac{|X^n_{\alpha,F}(s)|}{|X^n_\alpha(s)|}-\frac{|X^n_{\alpha,F}(t)|}{|X^n_\alpha(t)|}\right| > \frac{\epsilon}{2}\right] < \frac{\epsilon}{2}$$

for then we take $\delta_2=min(\delta_0,\delta_0')$. This is where our 'cutoff' of the process comes into play. At least for large enough $n$, we know that there will always be a large number of particles around. Observe that

$$\frac{|X^n_{\alpha,F}(s)|}{|X^n_\alpha(s)|}-\frac{|X^n_{\alpha,F}(t)|}{|X^n_\alpha(t)|}=\frac{1}{|X^n_\alpha(s)|}\Big(|X^n_{\alpha,F}(s)|-|X^n_{\alpha,F}(t)|\Big)$$

$$+\frac{1}{|X^n_\alpha(s)|}\cdot\frac{|X^n_{\alpha,F}(t)|}{|X^n_\alpha(t)|}\Big(|X^n_\alpha(s)|-|X^n_\alpha(t)|\Big)$$

and thus certainly

$$\left|\frac{|X^n_{\alpha,F}(s)|}{|X^n_\alpha(s)|}-\frac{|X^n_{\alpha,F}(t)|}{|X^n_\alpha(t)|}\right|\le\frac{1}{|X^n_\alpha(s)|}\Big(\big||X^n_{\alpha,F}(s)|-|X^n_{\alpha,F}(t)|\big|+\big||X^n_\alpha(s)|-|X^n_\alpha(t)|\big|\Big)$$

The lower bound on $|X^n_\alpha|$ makes this result an easy consequence of the

proof of Lemma 4.4.

We now deduce that $\{X^n\}_{n\geq 1}$ is relatively compact in the weak topology. Take a monotone decreasing sequence $\{\alpha_m\}_{m\geq 1}$ of rational numbers with $\alpha_m \downarrow 0$ as $m\to\infty$. We have already proved that $\{X^n_{\alpha_1}\}_{n\geq 1}$ has a convergent subsequence denoted $\{X^{n_i}_{\alpha_1}\}_{i\geq 1}$, say, with limit point $X^\infty_{\alpha_1}$. The corresponding subsequence $\{X^{n_i}_{\alpha_2}\}_{n\geq 1}$ of $\{X^n_{\alpha_2}\}_{n\geq 1}$ also has a convergent subsequence with limit point $X^\infty_{\alpha_2}$, say. From their construction it is clear that $X^\infty_{\alpha_1}(t)$, $X^\infty_{\alpha_2}(t)$ coincide for $t<\tau_{\alpha_1}$. Proceeding in this way we obtain a sequence of processes $\{X^\infty_{\alpha_m}\}_{m\geq 1}$ which are consistent in the above way and thus clearly have a unique limit point. We should like to see that this limit point is a weak limit point of the processes $\{X^n(t)\}_{n\geq 1}$. The only way in which this could fail would be if the processes $X^n$ were able to become arbitrarily small and then recover with strictly positive probability. That this is impossible is an easy consequence of the proof of Lemma 4.4, for observe that for $t>\tau_\alpha$

$$\mathbb{P}[|X^n(t)|>\alpha^{1/3}+\alpha]\leq \alpha^{1/3}\bar{\sigma}\bar{V}|t-\tau_\alpha|\leq \alpha^{1/3}\bar{\sigma}\bar{V}t$$

We have now proved the following

**Theorem 4.6** Suppose that $\{\nu_n\}_{n\geq 1}$ is a weakly convergent sequence of elements of $\mathcal{M}_0(\mathbb{R}^d)$ satisfying the conditions (4.1) and that $X^n(0)=\nu_n\ \forall n$. Then $\{X^n(\cdot)\}_{n\geq 1}$ is relatively compact in the weak topology.

§4.3 The weak limit points : A martingale characterization

In this section we identify the weak limit of a convergent subsequence of $\{X^n\}_{n\geq 1}$ (the subsequence also being denoted $\{X^n\}_{n\geq 1}$) as the unique solution to a martingale problem.

Recall that the distribution of $X^n$ is a solution to the $\mathcal{L}_n$-martingale problem over $D([0,\infty),E_n)$. In fact by assigning zero probability to those paths which are not $E_n$-valued, we may think of all these martingale problems as being over $D([0,\infty),\mathcal{M}_0(\mathbb{R}^d))$ or indeed any larger path space. The advantage of the martingale characterization is that if we can identify the limiting value of the operators $\mathcal{L}_n$ as an operator $\mathcal{L}$, say, acting on some appropriate class of functions, then we can deduce that the distribution of any weak limit point of $\{X^n(\cdot)\}_{n\geq 1}$ solves the martingale problem for $\mathcal{L}$ (acting on these functions and on this larger path space). If we can then prove uniqueness of the solution to this martingale problem, then we will have actually proved that $\{X^n(\cdot)\}_{n\geq 1}$ converges (weakly) and furthermore we will have a characterization of the limit. This is the strategy which we shall adopt.

Let us first then identify the operator $\mathcal{L}$. Suppose that $\phi\varepsilon C_b^+(\mathbb{R}^d)\cap\mathfrak{D}(L)$. Then at least for large enough $n$, if we set $f_n=(1-\phi/n)$, we will have $exp(<nlogf_n,\cdot>)\varepsilon\mathfrak{D}(\mathcal{L}_n)$. Notice that

$$\exp(<\text{nlogf}_n,\mu>)\rightarrow exp(-<\phi,\mu>) \qquad \text{as } n\rightarrow\infty$$

Consider then

$$\mathcal{L}_n(exp(<nlogf_n,\mu>))=\left\langle \frac{nLf_n+n^2(\Phi(f_n)-f_n)}{f_n}\ ,\ \mu\right\rangle\cdot exp(<nlogf_n,\mu>)$$

$$= -\left\langle \frac{L\phi + F_n(\phi)}{1-\phi/n} \, , \, \mu \right\rangle \cdot exp(<nlogf_n, \mu>)$$

where

$$F_n(\phi) = n^2 V(1-\Phi(1-\phi/n)-\phi/n)$$

Expanding $F_n(\phi)$ as a Taylor series in $\phi$, with integral form of the remainder yields

$$F_n(\phi) = -\phi^2 V \int_0^1 (1-u)\left[\frac{\partial^2 \Phi}{\partial s^2}\right]_{(1-\phi u/n)} du$$

where we have suppressed the dependence of both sides on $x$. Letting $n \to \infty$ we see that

$$F_n(\phi) \to -b\phi^2$$

where

$$b = \frac{1}{2} V\left[\frac{\partial^2}{\partial s^2}\Phi(s,x)\right]_{s=1}$$

By our assumptions on p.15, $b$ is a bounded continuous function.

Consider then $\mathcal{L}$ defined on $\mathcal{D} = \{exp(-<\phi,\mu>) : \phi \varepsilon C_b^+(\mathbb{R}^d)\}$ by

$$\mathcal{L}(exp(-<\phi,\mu>)) = <-L\phi + b\phi^2, \mu> exp(-<\phi,\mu>)$$

It is clear from above that for each function $F \varepsilon \mathcal{D}$, there exist functions $F_n \varepsilon \mathcal{D}(\mathcal{L}_n)$ such that $F_n \to F$ and $\mathcal{L}_n F_n \to \mathcal{L} F$ as $n \to \infty$.

We now know that the distribution of any limit point of $\{X^n(\cdot)\}_{n \geq 1}$ solves the

martingale problem for $(\mathcal{L},\nu)$ over $D([0,\infty),\mathcal{M}_0(\mathbb{R}^d))$. It remains to show that this uniquely determines the distribution of the limit point. To do this we employ 'the method of duality'. The version of duality which we use was appled in this way ('though in slightly greater generality) by Dawson&Kurtz (1982). In this simple case, the proof is straightforward and is given in Appendix C.

Theorem C.1 (The method of duality)

Let $E_1$, $E_2$ be metric spaces and suppose that $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$ are probability distributions on the space of càdlàg functions from $[0,\infty)$ to $E_1$, $E_2$ respectively. (These spaces are endowed with the natural $\sigma$-fields c.f. p.21.) Take two functions $f$, $g$ on $E_1\times E_2$ for which the following are true:

(i) $$\{\Big(f(\cdot,y),g(\cdot,y)\Big) : y\varepsilon E_2\}\subseteq C(E_1)\times C(E_1)$$

(ii) $$\{\Big(f(x,\cdot),g(x,\cdot)\Big) : x\varepsilon E_1\}\subseteq C(E_2)\times C(E_2)$$

(iii) For $y\varepsilon E_2$

$$f(X(t),y)-\int_0^t g(X(s),y)ds \text{ is a } \hat{\mathbb{P}}^{X(0)}\text{-martingale}$$

(iv) For $x\varepsilon E_1$

$$f(x,Y(t))-\int_0^t g(x,Y(s))ds \text{ is a } \tilde{\mathbb{P}}^{Y(0)}\text{-martingale}$$

(v) $f$, $g$ are bounded.

Then

(4.3) $$\hat{\mathbb{E}}^{X(0)}[f\Big(X(t),Y(0)\Big)]=\tilde{\mathbb{E}}^{Y(0)}[f\Big(X(0),Y(t)\Big)]$$

Consider then the functions $\{f(\phi,\mu)=exp(-<\phi,\mu>) : \phi \varepsilon C_b^+(\mathbb{R}^d),\ \mu \varepsilon \mathcal{M}_0(\mathbb{R}^d)\}$, and for each such $f$, set $g(\phi,\mu)=\mathcal{L}f(\phi,\mu)$. In place of $X(\cdot)$, we take one of our measure-valued limit processes whose distribution is a solution to the martingale problem for $\mathcal{L}$, and in place of $Y(\cdot)$, we take the deterministic flow, $u(\cdot)$ (if it exists), given by

$$\frac{\partial u}{\partial t}=Lu-bu^2 \qquad u(0)\varepsilon C_b^+(\mathbb{R}^d)\cap \mathcal{D}(L) \tag{4.4}$$

If we can show existence and uniqueness of the solution to equation (4.4), then from (4.3) we can deduce uniqueness of the solution to the martingale problem for $\mathcal{L}$. The second order non-linear partial differential equation (4.4) is called the *evolution equation* of the measure-valued process with generator $\mathcal{L}$. In practice we actually consider the equivalent integral equation—known as the *mild* form of the evolution equation

$$u(t)=T_t u(0)-\int_0^t T_{t-s}(b\cdot u(s)^2)ds \qquad u(0)=\phi \varepsilon C_b^+(\mathbb{R}^d) \tag{4.5}$$

where $\{T_t\}_{t\geq 0}$ is the semigroup of linear operators with generator $L$—i.e. the expectation semigroup of the diffusion part of the process. For the remainder of this chapter we will assume existence and uniqueness of the solution $u(\cdot)$ of (4.5) for $u(0)\varepsilon C_b^+(\mathbb{R}^d)$. Uniqueness follows easily from Picard's Theorem and we defer the proof of existence to Chapter 5. It will be clear there that if $u(0)\varepsilon \mathcal{D}(L)$, then $u(t)$ also solves equation (4.4), and moreover the solution to (4.5) depends (weakly) continuously on the initial data $\phi$ so that we can extend our martingale characterization to all $\phi \varepsilon C_b^+(\mathbb{R}^d)$.

Remark 1 (Boundary conditions)

Suppose that we wish to construct the measure-valued diffusion approximation to critical branching processes on some domain $G \subseteq \mathbb{R}^d$, then on hitting the boundary $\partial G$ of $G$ the particles will be subject to some sort of boundary condition. This will be reflected in some sort of boundary condition on functions in $\mathcal{D}(L)$ and in turn this will be translated into a boundary condition on the evolution equation (4.4). For example, if particles are reflected in direction $\gamma$ on hitting the boundary, then the measure-valued process will be characterized by the evolution equation (4.4), with the additional condition

$$\frac{\partial u}{\partial \gamma}\Big|_{\partial G} = 0$$

Remark 2 (Continuity)

Notice that in the proof of Proposition 4.3 we made no requirement on the choice of partition $\{t_i\}$ of $[0,T]$. It follows that $X(t)$ has continuous sample paths.

§4.4 Infinite initial measures

The measure-valued diffusions which we have constructed so far have all taken their initial values among the set of finite measures of compact support. In this section we indicate how to increase our repertoire to include all finite and some infinite initial measures.

In order to achieve a sensible definition of a measure-valued process started from

an infinite measure we should like to avoid the possibility that an infinite amount of mass might 'cascade in' from infinity in a finite time, and so we do not take arbitrary initial values. Instead, we consider measure-valued processes started from classes of 'tempered' measures analogous to those introduced (for some special cases) by Iscoe, Iscoe (1986b). Those details which we omit in constructing these spaces are entirely analogous to those in Iscoe's paper, 'though our construction of the process differs from his in appealing to Theorem (4.1) rather than proving estimates on our spatial diffusion semigroup and appealing to the general theory of Markov processes.

Recall that each of our measure-valued processes is multiplicative; that is the distribution of the process with initial value $\mu_1+\mu_2$ is the distribution of the sum of two independent copies of the process with initial values $\mu_1$, $\mu_2$ respectively. A natural way then to try to construct the measure-valued process started from an initial measure, $\mu$, with infinite support would be to decompose $\mu$ into a countable number of finite measures, $\{\mu_i\}$ say, of finite (disjoint) supports. We then try to construct the measure-valued process started from $\mu$ as the limit as $n\rightarrow\infty$ of the processes $X^{(n)}$, where $X^{(n)}(0)$ is given by

$$\mu^{(n)}=\sum_{i=1}^{n}\mu_i$$

From our Laplace transform characterization of the measure-valued process started from $\mu_i$, the measure-valued process started from $\mu$ (if it exists) will satisfy

$$\mathbb{E}_{\mu}[exp(-<\phi,\eta(t)>)]=\prod_{i=1}^{\infty}\mathbb{E}_{\mu_i}[exp(-<\phi,\eta(t)>)]$$

$$=\prod_{i=1}^{\infty}exp(-<u(t),\mu_i>)$$

$$= exp(-<u(t),\mu>)$$

where $u$ satisfies the usual evolution equation with $u(0)=\phi$. Thus to avoid the problem of 'cascading mass' we should choose $\mu$ in such a way that

$$\int u(t)d\mu < \infty \qquad\qquad \forall t \geq 0,\ \forall \phi \varepsilon C_K(\mathbb{R}^d)$$

Unfortunately, $u(t)$ is difficult to write down explicitly. However, we can certainly find an upper bound. For example, it is clear from (4.5) that $u(t) \leq T_t u(0)$ where $\{T_t\}_{t \geq 0}$ is the semigroup of linear operators corresponding to the spatial diffusion part of the process. By our assumptions on $L$, $T_t u(0)$ is strictly positive on $\mathbb{R}^d$ for each $t>0$. Further, if $\phi_1$, $\phi_2 \varepsilon C_K(\mathbb{R}^d)$, then, modulo a constant, $T_t\phi_1(x)$ and $T_t\phi_2(x)$ exhibit the same asymptotic behaviour (as a consequence of Assumption (v) on p.15). It is natural then to fix $\phi_0 \varepsilon C_K^+(\mathbb{R}^d)$ and topologize

$$\mathcal{M}_t = \{\mu \varepsilon \mathcal{M}(\mathbb{R}^d) : \int T_t \phi d\mu < \infty \ \forall \phi \varepsilon C_K(\mathbb{R}^d)\}$$

by the smallest topology making the maps

$$\{\mu \mapsto \int \phi d\mu,\ \ \phi \varepsilon C_K(\mathbb{R}^d) \cup T_t\phi_0\}$$

continuous. We should like to take an initial value

$$\mu \varepsilon \bigcap_{s \epsilon (0,\infty)} \mathcal{M}_s$$

and show that the corresponding processes $\{X^{(n)}\}_{n\geq 1}$ converge in $\mathcal{M}_t$ for each $t>0$.

Remark Notice in particular that if $\mu$ is an invariant measure for the spatial diffusion part of our process, then $\mu \varepsilon \mathcal{M}_t$ for all $t$.

The difficulty is that $\mathcal{M}_t$ is not complete in the given topology and so we must alter our setting slightly. Let $\hat{\mathbb{R}}^d=\mathbb{R}^d\cup\{\tau\}$ denote the one-point compactification of $\mathbb{R}^d$, and let $\mathcal{M}(\mathbb{R}^d)$ be the space of positive measures on $\hat{\mathbb{R}}^d$. We extend $\mu \varepsilon \mathcal{M}(\mathbb{R}^d)$ to $\hat{\mathbb{R}}^d$ by $\mu(\{\tau\})=0$. Write

$$\hat{C}_t=\left\{\phi \varepsilon C(\mathbb{R}^d) : \lim_{|x|\to\infty}\frac{|\phi(x)|}{|T_t\phi_0(x)|}=c \varepsilon [0,\infty) \text{ extended to } \hat{\mathbb{R}}^d \text{ by } \phi(\tau)=c\right\}$$

The space

$$\hat{\mathcal{M}}_t=\left\{\mu \varepsilon \mathcal{M}(\hat{\mathbb{R}}^d) : <T_t\phi_0,\mu>=\int T_t\phi_0 d\mu+\mu(\tau) < \infty\right\}$$

is topologized by the weakest topology making the maps $\{\mu\mapsto<\phi,\mu>: \phi \varepsilon C_K(\mathbb{R}^d)\cup T_t\phi_0\}$ continuous [For $\phi \varepsilon \hat{C}_t$, $\mu \varepsilon \hat{\mathcal{M}}_t$

$$<\phi,\mu>=\int\phi_1 d\mu_1+\phi(\tau)\mu(\tau)$$

where $\phi_1$, $\mu_1$ are respectively the restrictions to $\mathbb{R}^d$ of $\phi$, $\mu$]. If $d_0$ is a metric for the vague topology on $\mathcal{M}(\mathbb{R}^d)$ (under which it is complete) then it is a routine matter to check that

$$d(\mu_1,\mu_2)=d_0(\mu_1|_{\mathbb{R}^d},\mu_2|_{\mathbb{R}^d})+|<T_t\phi_0,\mu_1-\mu_2>|$$

defines a metric on $\hat{\mathcal{M}}_t$ under which it is complete and separable.

Considering $X^{(n)}$ as an $\hat{\mathcal{M}}_t$-valued process, it is easy now to verify the conditions of Theorem (4.1). Condition (b) requires no checking since the processes are all continuous. Condition (a) follows from the fact that $\{\mu \varepsilon \hat{\mathcal{M}}_t : <T_t\phi_0,\mu> \leq k\}$ is compact for each fixed $k \geq 0$.

The problem now is to show that if we denote a limit point of $\{X^{(n)}\}_{n\geq 1}$ by $X$, then $X_s(\tau)=0$ a.s. for all $s$. Let $\{\rho_n\}_{n=1}^{\infty}$ be a sequence of $C_K(\mathbb{R}^d)$ valued functions with $0 \leq \rho_n \leq 1$ and such that $\rho_n \uparrow 1$ as $n \to \infty$. Also, set $\sigma_n = 1-\rho_n$. Then

$$\mathbb{P}[X_s(\tau)>0] = \mathbb{E}[<\chi_{\{\tau\}}, X(s)>]$$

$$\leq \lim_{n\to\infty} \mathbb{E}[<\sigma_n T_t\phi_0, X(s)>] \quad \text{(by Fatou's Lemma)}$$

$$= \lim_{n\to\infty} (<u(s,\sigma_n T_t\phi_0),\mu>)$$

$$\leq \lim_{n\to\infty} (<T_s(\sigma_n T_t\phi_0),\mu>)$$

$$=0 \qquad \text{(by the dominated convergence theorem)}$$

Thus, $X$ is actually $\mathcal{M}(\mathbb{R}^d)$-valued.

It is also easy to see that (with probability 1) if $\mu \varepsilon \bigcap \mathcal{M}_s$, then $X \varepsilon \bigcap \mathcal{M}_s$. For such $\mu$, we have now shown the existence of an $\mathcal{M}(\mathbb{R}^d)$-valued process $X(t)$, characterized via its Laplace functionals

$$\mathbb{E}_\mu[exp(-<\phi,X(t)>)]=exp(-<u(t),\mu>)$$

where $u$ solves the usual evolution equation and $u(0)=\phi\varepsilon\bigcap C_s$. It is a simple matter now to characterize $X$ as that $\mathcal{M}_b(\mathbb{R}^d)$-valued process whose distribution is the unique solution to the martingale problem for $\mathcal{L}$ over $D([0,\infty),\bigcap\mathcal{M}_t)$.

## §4.5 Homeomorphisms of $\mathbb{R}^d$

We close this chapter with a simple observation which will later prove to be a useful device in reducing the study of our measure-valued processes with reflecting/absorbing boundary conditions to certain special cases.

Suppose that $\psi$ is a homeomorphism, $\psi:\mathbb{R}^d\to\mathbb{R}^d$, then this induces a natural map $\mathcal{M}_b(\mathbb{R}^d)\to\mathcal{M}_b(\mathbb{R}^d)$ via

$$\mu\mapsto\mu_\psi$$

$$\mu_\psi(B)=\mu(\psi^{-1}(B)) \qquad B\varepsilon\mathcal{B}(\mathbb{R}^d)$$

($\mathcal{B}(\mathbb{R}^d)$ denotes the Borel sets of $\mathbb{R}^d$). Thus we can regard $\psi$ as a transformation of measure-valued processes. Since $\psi$ maps the paths of a single particle diffusion with infinitesimal generator $L$ on $\mathbb{R}^d$ to the paths of the diffusion with generator $\hat{L}$ given by

$$L[f\circ\psi](x)=\hat{L}[f](\psi(x)) \qquad x\varepsilon\mathbb{R}^d$$

without time change (see Csink&Øksendal (1983) for a more general result), it is easy

to see that the measure-valued process with evolution equation (4.4) will be transformed into the measure-valued process with evolution equation

$$\frac{\partial u}{\partial t} = \hat{L}u - \hat{b}u^2$$

where $$\hat{b}(x) = b(\psi^{-1}(x))$$

**Remark** Of course, we could state this in a great deal more generality—there is no reason why we shouldn't take any continuous $\psi$ of the form considered by Csink&Øksendal.

## Chapter 5. A Solution to the Evolution Equation

In principal, the characterization of the measure-valued critical branching diffusions which we obtained in Chapter 4 should tell us everything we wish to know about the distribution of the processes. The difficulty is that the evolution equation is a second order non-linear partial differential equation, and we have no simple techniques for finding closed solutions to such equations. However, as we show in §5.1, it is possible to find a series solution to the evolution equation using a simple iteration technique. This approach has the added bonus of yielding expressions for the moments of the measure-valued diffusions. The remainder of this chapter is devoted to establishing the asymptotic behaviour of these coefficients as $t \to \infty$. Although we consider several disjoint cases, the techniques used are closely related to each other.

If the spatial diffusion part of our process is transient, then the coefficients of the series solution remain bounded for all times $t$. However, if the spatial diffusion part is recurrent, then they grow to infinity as $t \to \infty$. In order to simplify the investigation of the asymptotic behaviour, at the beginning of §5.2 we introduce the assumption that the branching mechanism, $b$, of the measure-valued processes should be constant (in fact we take this constant to be 1). Chapter 8 includes a brief discussion of the effects of dropping this assumption. In §5.2 we use Tauberian theory to obtain asymptotic expressions for the coefficients of the series in the case when the spatial diffusion part of our process is positive recurrent. These expressions will be a fundamental tool in Chapter 7. In §5.3 we consider the more difficult case of a null recurrent diffusion. Here, for the purposes of Chapter 6, we will be interested in the value of the integral of the coefficient with respect to the invariant measure $\mu$ of the

spatial diffusion given by Assumption (v) on p.15. We will no longer be able to find the exact asymptotic behaviour of the coefficients, but nonetheless we can find an upper bound on their rate of growth and we can find a lower bound on the rate of growth of the second order term.

## §5.1 A Wild sum

Recall from Chapter 4, that we may consider the so-called 'mild' form of the evolution equation.

$$u(t) = T_t u(0) - \int_0^t T_{t-s}\Big(b \cdot u(s)^2\Big) ds \tag{5.1}$$

where $u(0) \varepsilon C_b^+(\mathbb{R}^d)$. Suppose then that $u(0)$ is of the form $\lambda\phi$ for some $\lambda > 0$. We proceed formally. Define a multiplication operator $\mathcal{I}$ by

$$\mathcal{I}(f \circ g)(t) = -\int_0^t T_{t-s}\Big(b \cdot f(s) \cdot g(s)\Big)\, ds$$

Equation (5.1) becomes

$$u(t) = T_t u(0) + \mathcal{I}(u \circ u)(t) \tag{5.2}$$

It is easy to prove that $\mathcal{I}$ is distributive, but not, in general, associative. We try to solve (5.2) iteratively. Define

$$u_0(t) = T_t u(0)$$

$$u_{r+1}(t)=T_t u(0)+\mathcal{I}(u_r \circ u_r)(t)$$

It is clear that if the series $\{u_r\}_{r\geq 0}$ converges to a limit $u$, then $u$ will solve (5.2). We expand the first few members of the sequence.

$$u_1=u_0+\mathcal{I}(u_0 \circ u_0)$$

$$u_2=u_0+\mathcal{I}\Big((u_0+\mathcal{I}(u_0 \circ u_0)) \circ (u_0+\mathcal{I}(u_0 \circ u_0))\Big)$$

$$=u_0+\mathcal{I}(u_0 \circ u_0)+\mathcal{I}(u_0+\mathcal{I}(u_0 \circ u_0))+\mathcal{I}(\mathcal{I}(u_0 \circ u_0) \circ u_0)+\mathcal{I}\Big(\mathcal{I}(u_0 \circ u_0) \circ \mathcal{I}(u_0 \circ u_0)\Big)$$

Continuing in this way, we see a pattern emerging. It is clear that there is a one-to-one correspondence between terms of the expansion of this type of $u_r$ and the set $\mathcal{T}_r$ of ordered binary rooted trees of order less than or equal to $2^r$. (The reader is referred to Appendix D for definitions and results related to such trees.) Indeed we may write

$$u_r(t)=\sum_{\tau \in \mathcal{T}_r} U_t^{\tau}(\lambda \phi)$$

where $\{U_t^{\tau}(\lambda \phi)\}$ are defined inductively as follows. If $\tau_e$ is the tree of order 1 and $\tau_1 \circ \tau_2$ denotes the tree formed by linking the roots of trees $\tau_1$ and $\tau_2$ to a common root (thus forming a tree of order $|\tau_1|+|\tau_2|$ ) then

$$U_t^{\tau_e}(\lambda \phi)=\lambda T_t \phi$$

$$U_t^{\tau 1 \circ \tau 2}(\lambda\phi) = \mathcal{I}\Big( U^{\tau 1}(\lambda\phi) \circ U^{\tau 2}(\lambda\phi) \Big)(t)$$

Formally then

$$u(t) = \sum_{\tau \in \mathcal{T}} U_t^{\tau}(\lambda\phi) = \sum_{\tau \in \mathcal{T}} \lambda^{|\tau|} U_t^{\tau}(\phi) \tag{5.3}$$

is a solution to (5.2) (the sum being over all ordered binary rooted trees).

Remark Sums of this type are often called Wild sums. Wild (1951) used this construction to produce a series solution of Boltzmann's equation.

We investigate the convergence of this power series. Let us begin by supposing that $\phi$ is of the form $\phi = \chi_K$ (the indicator function of the set $K$) for some compact set $K \subseteq \mathbb{R}^d$. Write T(t) for the maximum expected sojourn time up to time t of a particle moving under the diffusion generated by $L$ in the set K; the maximum being taken over all initial positions of the particle. Then

$$U_t^{\tau e \circ \tau e} = \int_0^t T_{t-s}\Big( b \cdot T_s(\lambda\chi_K) \cdot T_s(\lambda\chi_K) \Big) ds$$

$$\leq \bar{b} \cdot \lambda^2 \cdot T(t) \cdot (T_t \chi_K)(x) \qquad \text{where } \bar{b} = \sup_x b(x)$$

Proceeding inductively, we have

$$|U_t^{\tau}(\lambda\chi_K)(x)| \leq \lambda^{|\tau|} \left( \bar{b}\ T(t) \right)^{|\tau|-1} (T_t \chi_K)(x)$$

Now, from Theorem (D.1), if $C(n)$ denotes the number of ordered binary rooted trees of order $n$, then the series

$$\sum_{n=1}^{\infty} C(n)\lambda^n\Big(\ \bar{b}T(t)\Big)^{n-1}$$

is convergent for $\lambda\bar{b}T(t) < \frac{1}{4}$. Thus, for $\lambda < \frac{1}{8\bar{b}T(t)}$, say

$$\sum_{\tau\varepsilon\mathcal{T}} U_t^{\tau}(\lambda\chi_K)(x)$$

is absolutely and uniformly (in both $\lambda$ and x) convergent to a uniformly bounded limit.

Obviously, we can find a corresponding bound on $\lambda$ for any $\phi\varepsilon C_K(\mathbb{R}^d)$ : the series solution is certainly absolutely and uniformly convergent for $\lambda\bar{b}\tilde{T}(t)\cdot\sup_x|\phi(x)|<\frac{1}{8}$, where $\tilde{T}(t)$ is the maximum expected sojourn time of a particle moving under the spatial diffusion in $supp\phi$.

**Remark** Notice then that if $L$ is the generator of a transient diffusion process on $\mathbb{R}^d$, then $T(t)$ is finite for all $t$, and so for sufficiently small $\lambda$, our series solution to the evolution equation is valid for all times $t$. If the diffusion is recurrent, then (as we would guess from above and will follow from §5.2 and §5.3) we cannot find sufficiently small $\lambda$ that the series solution should be valid for all time. However, as we shall see, this doesn't render the solution useless even for analysing asymptotic behaviour of the measure-valued processes.

An important aspect of the series solution is that repeatedly differentiating our expression for $\mathbb{E}_{X(0)}[exp(-<\lambda\phi,X(t)>)]$ with respect to $\lambda$ and evaluating at $\lambda=0$ yields expressions for the moments of the measure-valued process in terms of the coefficients of the power series. Moreover, it is no more difficult to establish joint distributions. For example, to calculate

$$\mathbb{E}[<\phi_1,X(t)>\cdot<\phi_2,X(t)>\cdot\ldots\cdot<\phi_n,X(t)>]$$

we consider the evolution equation with initial condition

$$u(0)=\lambda_1\phi_1+\lambda_2\phi_2+\ldots+\lambda_n\phi_n$$

and differentiate in turn in $\lambda_1,\ldots,\lambda_n$ and evaluate at $\lambda_1=\lambda_2=\ldots=\lambda_n=0$.

§5.2 Asymptotic estimates : Positive recurrent spatial diffusion

In theory we could now write down an expression for any moment of the measure-valued branching process in terms of the linear semigroup corresponding to its spatial diffusion part. In practice, of course, this would be a tedious exercise. Instead we satisfy ourselves with establishing some asymptotic estimates for the coefficients of our power series. We are interested in the case where the spatial diffusion part of the process is recurrent. In order to simplify our calculations,

For the remainder of the thesis we assume the branching mechanism of the measure-valued process to be constant.

Indeed we may as well set $b\equiv 1$.

Suppose that $a$ is the density of a positive measure $A$ concentrated on $(0,\infty)$. We write $A(x)$ for $A\{(0,x]\}$ and use $\omega(\lambda)$ to denote the Laplace transform of $A$.

$$\omega(\lambda)=\int_0^\infty exp(-\lambda x)A(dx)=\int_0^\infty exp(-\lambda x)\,a(x)\,dx$$

Our calculations rest on the application of the following results from Tauberian theory (see, for example, Feller (1966) Vol. 2, ChapterXIII). Recall once more that a function $l(t)$ on $(0,\infty)$ is said to be slowly varying at infinity if for every fixed $x$,

$$\frac{l(tx)}{l(t)}\to 1 \qquad \text{as } t\to\infty$$

Theorem 5.1 If $l$ is slowly varying at infinity and $0\le\rho<\infty$, then each of the relations

$$\omega(\tau)\sim\frac{1}{\tau^\rho}l\left(\frac{1}{\tau}\right) \qquad \tau\to 0 \tag{5.4}$$

$$A(t)\sim\frac{1}{\Gamma(\rho+1)}t^\rho l(t) \qquad t\to\infty \tag{5.5}$$

implies the other (where $\Gamma$ is Euler's gamma fuction).

Theorem 5.2 Let $0<\rho<\infty$. If $a$ is ultimately monotone then as $\lambda\to 0$ and $x\to\infty$, respectively,

$$\omega(\lambda)\sim\frac{1}{\lambda^\rho}l\left(\frac{1}{\lambda}\right) \quad iff \quad a(x)\sim\frac{1}{\Gamma(\rho)}x^{\rho-1}l(x) \tag{5.6}$$

Corollary 5.3 Suppose that $a$ is ultimately monotone. If (5.5) holds with $\rho>0$ then

$$a(x)\sim\frac{\rho A(x)}{x} \qquad x\to\infty \tag{5.7}$$

Let us then consider the coefficients of the series solution to the evolution equation in the case where the spatial diffusion part of the process is a positive recurrent diffusion. From the assumptions on p.15 we then know that

$$\frac{1}{t}\int_0^t T_s\phi\,ds\to\int\phi\,d\mu \qquad \text{as } t\to\infty$$

where $\mu$ is a non-trivial invariant measure of the spatial diffusion. Corollary 5.3 then implies that

$$T_t\phi\sim\int\phi\,d\mu \qquad \text{as } t\to\infty \tag{5.8}$$

Observe also that if we define the resolvent $R_\lambda$ corresponding to the semigroup $\{T_t\}_{t\geq 0}$ in the usual way, then from Theorem 5.1

$$\lambda R_\lambda\phi=\lambda\int_0^\infty exp(-\lambda s)\,T_s\phi\,ds\sim\Gamma(2)\int\phi\,d\mu=\int\phi\,d\mu \qquad \text{as } \lambda\to 0 \tag{5.9}$$

Our first aim is to estimate $\int U_t^{\tau^s}\phi\,d\mu$. Now an application of Theorem 5.2 and (5.8) yields

$$\int_0^\infty exp(-\lambda s)\, T_s\phi \cdot T_s\phi\, ds \sim \Gamma(1)\frac{1}{\lambda}\Big(\int \phi\, d\mu\Big)^2 \qquad \text{as } \lambda \to 0$$

and from an application of Theorem 5.1

$$\int_0^t T_s\phi \cdot T_s\phi\, ds \sim t \cdot \frac{\Gamma(1)}{\Gamma(2)}\Big(\int \phi\, d\mu\Big)^2 \qquad \text{as } t \to \infty$$

We integrate w.r.t. $\mu$ to obtain

$$\iint_0^t T_s\phi \cdot T_s\phi\, ds d\mu \sim t \cdot \frac{\Gamma(1)}{\Gamma(2)}\Big(\int \phi\, d\mu\Big)^2 \qquad \text{as } t \to \infty \tag{5.10}$$

(Strictly speaking we justify this by dividing by $t$ and applying the Dominated Convergence Theorem.)

We now use (5.9) to arrive at an asymptotic expression for $U_t^\tau \phi$. Observe then that

$$\int exp(-\lambda t)\int_0^t T_{t-s}(T_s\phi \cdot T_s\phi)\, ds dt = \int_0^\infty \int_0^\infty exp(-\lambda r)\, exp(-\lambda s)\, T_r(T_s\phi \cdot T_s\phi)\, ds dr$$

$$= R_\lambda\Big(\int_0^\infty exp(-\lambda s)\, T_s\phi \cdot T_s\phi)\, ds\Big)$$

$$\sim \frac{1}{\lambda}\Big(\int_0^\infty exp(-\lambda s)\Big(\int T_s\phi \cdot T_s\phi\, d\mu\Big) ds\Big) \qquad \text{as } \lambda \to 0$$

$$\sim \frac{1}{\lambda} \cdot \frac{1}{\lambda} \cdot \frac{\Gamma(1)}{\Gamma(2)}\Big(\int \phi\, d\mu\Big)^2 \Gamma(2) \qquad \text{as } \lambda \to 0$$

using (5.10) and an application of Theorem 5.1. Theorem 5.2 now yields

(5.11) $$\int_0^t T_{t-s}(T_s\phi \cdot T_s\phi)ds \sim t \cdot \frac{\Gamma(1)}{\Gamma(2)}\left(\int \phi d\mu\right)^2 = t \cdot \left(\int \phi d\mu\right)^2$$

i.e. $$\int_0^t U_t^{\tau s}\phi ds \sim -t \cdot \left(\int \phi d\mu\right)^2$$

We can now repeat this procedure to obtain higher order terms. We are going to apply an inductive argument. Suppose then that for two trees $\tau_1$ and $\tau_2$, we have established that

$$U_t^{\tau i}\phi \sim C_i t^{|\tau_i|-1}\left(\int \phi d\mu\right)^{|\tau_i|} \qquad \text{as } t \to \infty$$

Write $\tau = \tau_1 \circ \tau_2$. Then in exactly the same way as above we begin by applying Theorem 5.2 to obtain

$$\int_0^\infty exp(-\lambda s)\, U_s^{\tau 1}\phi \cdot U_s^{\tau 2}\phi ds \sim C_1 C_2 \left(\frac{1}{\lambda}\right)^{|\tau|-1}\left(\int \phi d\mu\right)^{|\tau|}\Gamma(|\tau|-1) \qquad \text{as } \lambda \to 0$$

and then by an application of Theorem 5.1

$$\int_0^t U_s^{\tau 1}\phi \cdot U_s^{\tau 2}\phi ds \sim C_1 C_2 t^{|\tau|-1}\left(\int \phi d\mu\right)^{|\tau|}\frac{\Gamma(|\tau|-1)}{\Gamma(|\tau|)} \qquad \text{as } t \to \infty$$

Applying the Dominated Convergence Theorem this time yields

$$\int\int_0^t U_s^{\tau 1}\phi \cdot U_s^{\tau 2}\phi ds d\mu \sim C_1 C_2 t^{|\tau|-1}\left(\int \phi d\mu\right)^{|\tau|}\frac{\Gamma(|\tau|-1)}{\Gamma(|\tau|)} \qquad \text{as } t \to \infty$$

and then

$$\int_0^\infty exp(-\lambda t)\, U_t^\tau \phi\, dt \sim -\frac{1}{\lambda}\cdot C_1 C_2 \Big(\frac{1}{\lambda}\Big)^{|\tau|-1}\Big(\int \phi\, d\mu\Big)^{|\tau|}\frac{\Gamma(|\tau|-1)}{\Gamma(|\tau|)}\cdot\Gamma(|\tau|) \quad \text{as } \lambda\to 0$$

and finally using Theorem 5.2

$$U_t^\tau \phi \sim -C_1 C_2 t^{|\tau|-1}\Big(\int \phi\, d\mu\Big)^{|\tau|}\frac{\Gamma(|\tau|-1)}{\Gamma(|\tau|)} = \frac{-1}{(|\tau|-1)} C_1 C_2 t^{|\tau|-1}\Big(\int \phi\, d\mu\Big)^{|\tau|} \quad \text{as } \lambda\to 0$$

Thus the asymptotic behaviour of two trees of the same order is, modulo a constant multiple, the same. Let us then establish the asymptotic behaviour of the sum, $u_n(t)$, of the coefficients of our series solution to the evolution equation which correspond to trees of order $n$. Now, there is only one tree of order one, and so we have already established the behaviour of $u_1(t)$. We continue inductively. Suppose that we have established that for all $m \leq n$

$$u_m(t) \sim (-t)^{m-1}\Big(\int \phi\, d\mu\Big)^m \qquad \text{as } t\to\infty$$

Every tree of order $n$ is built up in a unique way from the two trees of lower order obtained by removing its root, and so from the calculation above it is easy to see that

$$u_n(t) \sim \sum_{p=1}^{n-1} \frac{(-t)^{n-1}}{(n-1)}\Big(\int \phi\, d\mu\Big)^n = (-t)^{n-1}\Big(\int \phi\, d\mu\Big)^n \qquad \text{as } t\to\infty$$

Our proof actually yields more than this. We also have that

(5.12) $$\int u_n(t)\, d\mu \sim (-t)^{n-1}\Big(\int \phi\, d\mu\Big)^n \qquad \text{as } t\to\infty$$

§5.3 Asymptotic estimates : Null recurrent spatial diffusion

We now turn to the case of a null recurrent spatial diffusion. As we shall see matters are considerably more difficult. In particular, we will no longer have the use of the Dominated Convergence Theorem, so that while we can still find upper bounds on the rate of growth of coefficients, exact evaluations would almost certainly require a precise knowledge of the transition probabilities of the spatial motion.

Let us begin with the important special case when the spatial diffusion is plane Brownian motion—in this case we have

$$\int_0^t T_s\phi\, ds \sim \frac{1}{2\pi} log t \qquad \text{as } t\to\infty$$

and

$$T_t\phi \sim \frac{1}{2\pi t}\cdot\int \phi\, d\mu \qquad \text{as } t\to\infty$$

where $\mu$ is now Lebesgue measure on $\mathbb{R}^2$ (we have taken the generator of Brownian motion to be $\frac{1}{2}\Delta$). Using the known form of the transition probabilities for Brownian motion, it is straightforward in this case to calculate the integral with respect to $\mu$ of the second order term in the series expansion

$$-\int U_t^{r_s}\phi\, d\mu = \int_0^t\int\int \frac{1}{4\pi s} exp\left(-\frac{|u-v|^2}{2s}\right)\phi(u)\phi(v)\mu(du)\mu(dv)ds$$

(see Dawson (1977)—his calculation is not affected by dropping the assumption that $\phi$ has compact support and for fixed $t$ it does not matter that we have not made his

assumption of transience!). Asymptotically then

$$-\int U_t^{\tau^s}\phi\,d\mu \sim \tfrac{1}{2}l(t)\Big(\int \phi\,d\mu\Big)^2 \qquad \text{as } t\to\infty$$

For a symmetric stable process of index $\alpha$, the factor of $1/2$ is replaced by one of $2^{-\alpha/2}$, so that one might conjecture that for all recurrent spatial diffusions we can find a constant $\beta$ with $0<\beta\leq 1$ such that

$$-\int U_t^{\tau^s}\phi\,d\mu \geq \beta l(t)\Big(\int \phi\,d\mu\Big)^2 \qquad \text{as } t\to\infty \tag{5.13}$$

One of our aims in this section is to prove that this is indeed the case.

An important difference between the asymptotic bounds which we obtain here and the estimates used in §5.1 to establish convergence of the series (5.3) is that the essential ingredient in these asymptotic expressions will not be expressed in terms of the sojourn time of a particle moving under the spatial diffusion in $suppu(0)$, but will instead depend on $\int u(0)\,d\mu$, where $\mu$ is the invariant measure of the spatial diffusion given by Assumption (v) on p.15. This means that we need not restrict our attention to $u(0)$ of compact support. This will be crucial in Chapter 6.

We will begin by establishing (5.13). We adopt a strategy closely related to that of §5.3. By assumption, there is a function $h(t)$ with $h(t)\to\infty$ as $t\to\infty$ for which

$$\frac{1}{h(t)}\int_0^t T_s\phi\,ds \to \int \phi\,d\mu \qquad \text{as } t\to\infty$$

where

$$h(t)\sim t^{\rho}l(t) \qquad \text{as } t\to\infty$$

for some slowly varying function $l$ and some $0\leq\rho<1$. Using Theorem 5.1 we see that this implies

$$R_\lambda\phi=\int_0^\infty exp(-\lambda s)\,T_s\phi\,ds\sim\left(\frac{1}{\lambda}\right)^{\rho}l\left(\frac{1}{\lambda}\right)\Gamma(\rho+1)\int\phi\,d\mu \qquad \text{as } \lambda\to 0 \tag{5.14}$$

(where $R_\lambda$is again the resolvent).

We divide the proof into two parts, $\rho>0$ and $\rho=0$.

Let us begin with the second order term in the case when $\rho>0$. In this case we can apply Corollary 5.3 and we see that

$$T_t\phi\sim\frac{\rho h(t)}{t}\int\phi\,d\mu \qquad \text{as } t\to\infty$$

Consider then $U_t^{Ts}\phi$. It is a trivial application of Jensen's inequality to see that

$$-U_t^{Ts}\phi\equiv\int_0^t T_{t-s}(T_s\phi\cdot T_s\phi)\,ds\geq t(T_t\phi)^2\sim\frac{\rho^2h(t)^2}{t}\left(\int\phi\,d\mu\right)^2 \qquad \text{as } t\to\infty \tag{5.15}$$

We can now reverse the procedure of (5.10) to (5.11). Applying then (in order) Theorem 5.2, (5.14) and Theorem 5.1 we deduce that

$$\int\int_0^t T_s\phi\cdot T_s\phi\,ds\,d\mu\geq\frac{\rho^2\Gamma(2\rho)}{(\Gamma(\rho+1))^2}\cdot h(t)\left(\int\phi\,d\mu\right)^2 \qquad \text{as } t\to\infty \tag{5.16}$$

It is well known (see e.g. Feller (1966) Vol.II Chapter 2) that

$$\frac{\rho^2\Gamma(2\rho)}{(\Gamma(\rho+1))^2}=\frac{\Gamma(2\rho)}{(\Gamma(\rho))^2}=\left(\int_0^1\theta^{\rho-1}(1-\theta)^{\rho-1}d\theta\right)^{-1}$$

and this is obviously monotone increasing in $\rho$. Thus (5.16) is of the required form.

In the case $\rho=0$, the estimate in (5.15) is too crude and we can no longer use Corollary 5.3. However, we are able to employ a very similar argument. Suppose then that we can show that for some $\beta>0$,

$$-\int_0^T U_t^{T\,s}\phi\,dt\sim\beta l(T)^2\left(\int\phi\,d\mu\right)^2 \qquad \text{as } T\to\infty \tag{5.17}$$

We again reverse the steps from (5.10) to (5.11) except that we must apply Theorem 5.1 in place of Theorem 5.2 to deduce that

$$-\int U_t^{T\,s}\phi\,d\mu\sim\beta l(t)\left(\int\phi\,d\mu\right)^2 \qquad \text{as } t\to\infty$$

which is of the required form. Our aim then is to prove (5.17). Rather than give a detailed general result, we shall indicate the technique by means of an example. Suppose then that $l(t)=\log t$. It is an easy exercise to prove that $\exists\epsilon,\epsilon'>0$ (both independent of $\phi$) for which,

$$\int_0^T\int_0^t T_s\phi\cdot T_s\phi\,ds\,dt>\epsilon(\log t)^2\left(\int\phi\,d\mu\right)^2 \qquad \text{as } t\to\infty$$

and setting $f(T)=T^2$,

$$\int_T^{f(T)} T_s\phi\cdot T_s\phi\,ds>\epsilon'\int_0^T T_s\phi\cdot T_s\phi\,ds \qquad \text{as } T\to\infty$$

Observe that

$$-\int_0^{f(T)} U_t^{\tau_s}\phi\,dt=\int_0^{f(T)}\int_0^{f(T)-u} T_u(T_s\phi\cdot T_s\phi)\,ds\,du$$

Now

$$\int_0^{f(T)}\int_0^{f(T)-u} T_u(T_s\phi\cdot T_s\phi)\,ds\,du\geq\int_0^{T} T_u\Big(\int_0^{f(T)-u} T_s\phi\cdot T_s\phi\,ds\Big)du$$

$$\geq\int_0^{T}\int_0^{f(T)-T}(T_{u+s}\phi)^2\,ds\,du$$

$$=\int_0^{T}\int_T^{f(T)} T_s\phi\cdot T_s\phi\,ds\,du$$

$$>\epsilon'\int_0^{T}\int_0^{T} T_s\phi\cdot T_s\phi\,ds\,du$$

$$>\epsilon\epsilon'(\log T)^2\Big(\int\phi\,d\mu\Big)^2 \qquad \text{as } T\to\infty$$

Now $\log(T^2)=2\log T$ implies the result (with $\beta=\epsilon\epsilon'/4$).

Let us then establish upper bounds for the rate of increase of the terms of the expansion. Now, from §5.1, we know that if $\phi=\chi_K$ (the indicator function of a compact set) then

$$\Big|\int U_t^{\tau_s}\chi_K\,d\mu\Big|\leq T(t)\Big(\int\chi_K\,d\mu\Big)$$

where $T(t)$ denotes the maximum expected sojourn time in the set $K$ of a particle moving under the spatial diffusion corresponding to $T_t$. Observe that

$$T(t) = \max_x \int_0^t T_s \chi_K(x)\, ds \sim h(t) \int \chi_K \, d\mu \qquad \text{as } t \to \infty$$

A density argument and linearity of $T_t$ then yield

$$\left| \int U_t^{\tau_s} \phi \, d\mu \right| \leq h(t) \left( \int \phi \, d\mu \right)^2 \qquad \text{as } t \to \infty, \forall \phi \, \varepsilon \, C_b(\mathbb{R}^d)$$

Finally, an induction identical to that of §5.1 leads to

$$\left| \int u_n(t) \, d\mu \right| \leq C(n) h(t)^{n-1} \left( \int \phi \, d\mu \right)^n \qquad \text{as } t \to \infty$$

where, as usual, $u_n$ denotes the sum of the coefficients of the series corresponding to trees of order $n$ and $C(n)$ denotes the number of trees of order $n$.

## Chapter 6. Asymptotic Behaviour of the Measure-valued Diffusions

In this chapter we investigate the asymptotic behaviour of the measure-valued diffusions constructed in Chapter 4. We begin in §6.1 with the simple case of a process with finite initial mass. In this case it is easily shown that the process suffers extinction; i.e. the total mass tends to zero as time increases to infinity. We therefore turn to infinite initial masses in search of interesting asymptotic behaviour. From the biological point of view, it is interesting to establish the existence or non-existence of non-trivial limiting distributions for this type of model. If such non-trivial limits do exist, then we should expect their domains of attraction to be dependent on the spatial diffusion part of the process. An obvious strategy then is to start one of these processes off from an invariant measure of the spatial diffusion and see if (at least for some values of the branching function $b$), the processes converge. In the special case where the spatial diffusion is Brownian motion or a symmetric stable process, this problem was investigated by Dawson (1977). In this case it is observed that there is a dichotomy between those processes whose spatial diffusion part is transient and those whose spatial diffusion part is recurrent. Perhaps surprisingly at first sight, the measure-valued diffusion has a non-trivial limiting distribution if the spatial diffusion is transient, but suffers local extinction—that is the expected amount of mass in any compact set tends to zero—if the spatial diffusion is recurrent. Intuitively, the explanation for this is simply that the transient process is constantly able to 'recharge' the finite set with mass which was initially residing at infinity. In §6.2 we extend Dawson's proof that if the spatial diffusion part of the measure-valued diffusion is transient, then the process started from a non-trivial invariant measure of

this diffusion has a non-trivial limiting distribution to processes with general transient spatial diffusions. In §6.3 we consider the case where the spatial diffusion is recurrent. Here Dawson's proof rests heavily on a precise knowledge of the transition probabilities of Brownian motion and the symmetric stable processes and so we are forced to adopt a different approach. The method which we employ mimics the proof in the transient case in that it uses the series expansion of the evolution equation, but since the range of initial values for which such a solution is valid decreases with time, we cannot find an exact asymptotic expression for the solution in this way and our argument rests on sublinearity and the estimates obtained in Chapter 5. It is no surprise that the dichotomy extends to this general case—and in fact we do not need to restrict ourselves to local extinction in the recurrent case; instead we show that if $\mu$ denotes an invariant measure of the spatial diffusion part of the process, then the integral of any $\mu$-integrable function with respect to the process tends to zero as $t \to \infty$. This result can be considered as the continuous analogue of the corresponding result for critical cluster fields (Kallenberg (1975)) where the clusters are replaced by random measures.

In §6.4 we briefly consider the corresponding results for processes on regions with a boundary.

### §6.1 Processes with finite initial measures

If we are considering finite measure-valued processes, then we may find the distribution of their total mass process by setting $u(0) \equiv 1$, in the evolution equation. By assumption we have that

$$b_{inf} = \inf_x b(x) < \infty$$

and so it is certainly the case that

$$u(t,x) \leq v(t,x) \qquad \forall t \geq 0,\ \forall x \in \mathbb{R}^d$$

where $v$ solves

$$\frac{\partial v}{\partial t} = Lv - b_{inf} v^2$$

$$v(0) = u(0)$$

Hence if $u_1$, $v_1$ denote the solutions to these evolution equations with $u(0) = 1 = v(0)$, then

$$u(t) \leq v(t) = \frac{1}{b_{inf} t + 1}$$

and then since the initial measure $\mu$ is assumed finite,

$$\int u(t)\, d\mu \to 0 \qquad \text{as } t \to \infty$$

i.e. the measure-valued process dies out.

### §6.2 Infinite initial measures : Transient spatial motion

Let us turn now to infinite initial measures. In Rosen (1975), it is established that for the Dawson process in $\mathbb{R}^3$, there is a positive lower bound on the total mass in a given

compact set—the bound being uniform in $t$ ('though he considers the evolution equation in a different context). We can therefore expect a non-trivial limiting distribution for this process. In Dawson (1977) this is confirmed. Furthermore, it is shown there that the result extends to the class of transient symmetric stable processes. In this section we prove the analogous result for measure-valued processes with general transient spatial diffusions (subject to the conditions of §3.1 and with the branching function $b\equiv 1$). The proof is closely related to that of Dawson. Here, however, in order to be consistent with our previous work we use Laplace transforms in place of his characteristic functionals. Most of the work was done in Chapter 5!

Theorem 6.1(a) Suppose that the spatial diffusion part of the measure-valued process with generator $\mathcal{L}$ is a transient diffusion on $\mathbb{R}^d$ with non-trivial invariant measure $\mu$. Then the $\mathcal{L}$-process with initial value $\mu$ has a non-trivial limiting distribution.

Remark Recall from §4.4 that the measure-valued process with this initial measure does exist!

The proof will be a simple consequence of the results of Chapter 5 and the following well-known result.

Theorem 6.2 (see Feller (1966) Vol.II, Chapter XIII, Theorem 2)
For $n=1, 2, \ldots$ let $F_n$ be a probability distribution with Laplace transform $\phi_n$.

(a) If $F_n \to F$ where $F$ is a possibly defective distribution with Laplace transform $\phi$ then $\phi_n(\lambda)\to\phi(\lambda)$ for $\lambda>0$.

(b) Conversely, if the sequence $\{\phi_n(\lambda)\}$ converges for each $\lambda>0$ to a limit $\phi(\lambda)$,

then $\phi(\lambda)$ is the transform of a possibly defective distribution $F$, and $F_n \to F$.

The limit $F$ is not defective *iff* $\phi(\lambda) \to 1$ as $\lambda \to 0$.

Proof of 6.1(a) Recall once more our Laplace transform characterization of the measure-valued process

$$\mathbb{E}_\mu[exp(-<\lambda\phi, X(t)>)] = exp(-<u_\lambda(t), \mu>)$$

where $u_\lambda(t)$ solves the usual evolution equation with $u(0) = \lambda\phi$.

In order to apply Theorem 6.2, we should like to see that $\int u_\lambda(t)\,d\mu$ converges to a finite limit as $t \to \infty$ for each $\lambda > 0$. But observe that this expression is monotone decreasing in $t$, and

$$0 \leq \int u_\lambda(t)\,d\mu \leq \int \lambda\phi\,d\mu$$

Therefore the integral does indeed converge to a bounded limit for each $\lambda$ and (by considering arbitrary subsequences $t_i$ of times) we may deduce from Theorem 6.2(b) the existence of a limiting distribution. We must establish that this limit is non-trivial.

Recall (from Chapter 5) that for sufficiently small $\lambda$, we can write down $u_\lambda(t)$ as a power series in $\lambda$ whose coefficients are expressed in terms of the linear semigroup corresponding to the spatial diffusion part of the process.

$$u_\lambda(t) = \sum_{\tau \in \mathcal{T}} \lambda^{|\tau|} U_t^\tau \phi$$

Let $\phi$ be the indicator function of a compact set $K \subseteq \mathbb{R}^d$, then if $T(t)$ denotes, as usual, the maximum expected sojourn time up to time $t$ of a particle in $K$, this series converges absolutely and uniformly for $\lambda < 1/(8T(t))$. Since the process is transient, $T(t) \leq T$ for some $T > 0$ (uniformly in $t$). Therefore we can integrate term by term to obtain an expression for $\int u_\lambda(t) d\mu$—this series also converges for $\lambda < 1/(8T)$. That the limiting distribution is non-trivial follows since the coefficients of the series are not all zero, so for at least some $\lambda$, the series has a non-trivial finite sum.

Since all the coefficients in the series are bounded, it is clear that this limiting distribution is not defective.

Remark In the special case of the Dawson process we can obviously say a lot more about the limiting distribution. For example, it is clear that it is spatially homogeneous.

### §6.3 Infinite initial measures : Recurrent spatial motion

Consider now the case where the spatial diffusion corresponding to $T_t$ is null recurrent (the case of positive recurrence being covered by §6.1). We will prove the following

Theorem 6.1(b) Suppose that the spatial diffusion part of the $\mathcal{L}$-superprocess is a recurrent diffusion on $\mathbb{R}^d$ with non-trivial invariant measure $\mu$. Then the $\mathcal{L}$-superprocess with initial value $\mu$ has no non-trivial limiting distribution.

Our proof requires a lemma.

Lemma 6.3 Suppose that $\beta>0$. Then we can find $\alpha\varepsilon(0,1)$ with

$$\frac{1}{\alpha}\left(\sqrt{1+\alpha}-\sqrt{1-\alpha}\right)<1+\frac{\alpha\beta}{16}$$

Proof Equivalently we seek $\alpha$ for which

$$2<\left(1+\frac{\alpha\beta}{16}\right)\left(\sqrt{1+\alpha}+\sqrt{1-\alpha}\right)$$

Notice then that the function on the right hand side is smooth on [0,1) and is equal to 2 at 0, with strictly positive first derivative at 0. The result follows.

Proof of 6.1(b) Again most of the work was done in Chapter 5. We argue by contradiction. Suppose that the process does converge to a non-trivial limit. Then, by Theorem 6.2(a), we must have that for any $\phi\varepsilon C_b(\mathbb{R}^d)$, $\int u(t)d\mu$ converges to a finite limit as $t\to\infty$ and furthermore this limit is non-trivial for at least some $\phi_0$. Let us choose such a $\phi_0$, and suppose that $T$ is chosen sufficiently large that

$$\left(1-\frac{\alpha\beta}{16}\right)\int u(T)d\mu<\lim_{t\to\infty}\int u(t)d\mu$$

where $\beta$ is chosen (as in (5.13)) in such a way that $(\forall\phi\varepsilon C_b^+(\mathbb{R}^d))$

$$\int U_t^{T^s}\phi\,d\mu\geq\beta h(t)\left(\int\phi\,d\mu\right)^2 \qquad \text{as } t\to\infty$$

and $\alpha$ is chosen as in Lemma 6.3. We show that there is a finite time $\tilde{T}$ at which

$$\int u(T+\tilde{T})\,d\mu < \left(1-\frac{\alpha\beta}{16}\right)\int u(T)\,d\mu$$

and, since $\int u(t)\,d\mu$ is monotone, thereby arrive at a contradiction.

Consider once again the series solution to the evolution equation. Since the spatial diffusion is recurrent, we have seen that we cannot find $\lambda>0$ for which the power series expansion for $u_\lambda(t)$ converges for all $t$. However, for fixed $t>0$, we do have convergence for $\lambda=\lambda(t)<(4h(t)\int\phi\,d\mu)^{-1}$. We set $\phi=u(T)$ in this series and for convenience of notation consider our time scale to have 'restarted' with $T=0$. We will exploit the sublinearity of the evolution equation. Observe then that for $\lambda\leq 1$,

$$u_1(t)\leq\frac{1}{\lambda}u_\lambda(t)$$

We set

$$\lambda(t)=\frac{\alpha}{4h(t)\int\phi\,d\mu} \qquad 0\leq\alpha<1-\epsilon$$

(where $\epsilon$ is a small positive constant). From Chapter 5, for this value of $\lambda$ we may express $\int u_{\lambda(t)}(t)\,d\mu/\lambda(t)$ as a series. (The rôle of $\epsilon$ is to ensure that we are strictly within the radius of convergence of the power series for $u_{\lambda(t)}(t)$, so that we may interchange the order of integration and summation.) Consider the first two terms

$$\frac{1}{\lambda(t)}\int U_t^{\tau^e}(\lambda(t)\phi)\,d\mu=\int\phi\,d\mu$$

$$\frac{1}{\lambda(t)}\int U_t^{\tau^s}(\lambda(t)\phi)\,d\mu\leq-\lambda(t)\beta h(t)\left(\int\phi\,d\mu\right)^2=-\frac{\alpha\beta}{4}\int\phi\,d\mu \qquad \text{as } t\to\infty$$

Thus, at least for sufficiently large $T$, the contribution to the sum from the first two terms is bounded above by

$$\left(1-\frac{\alpha\beta}{2\cdot 4}\right)\int\phi\,d\mu$$

Consider then the sum of the remaining *positive* terms of the series. These come from trees of odd order and from Chapter 5 we have that

$$\left|\int U_t^\tau(\lambda(t)\phi)\,d\mu\right|\leq h(t)^{|\tau|-1}\left(\lambda(t)\int\phi\,d\mu\right)^{|\tau|} \qquad \text{as } t\to\infty$$

From Appendix D, if $C(n)$ denotes the number of binary rooted trees of order $n$, then

$$\sum_{n=1}^{\infty}C(n)\theta^n=\frac{1}{2}\left(1-\sqrt{1-4\theta}\right)$$

Thus the sum of the positive terms of order greater than one in our expansion of $\int u_{\lambda(t)}\,d\mu/\lambda(t)$ is bounded above (for large enough $t$) by

$$\left(\frac{1}{4}\left(\sqrt{1+\alpha}-\sqrt{1-\alpha}\right)-\frac{\alpha}{4}\right)\int\phi\,\mathrm{d}\mu\cdot\frac{4}{\alpha}$$

But by our choice of $\alpha$, this is less than

$$\frac{\alpha\beta}{16}\int\phi\,d\mu$$

and this leads us to the required contradiction.

Remark The proof above uses the crudest of estimates and so almost certainly disguises the rather rapid rate of decay of the process at least in a compact set. We will return to this rate of decay again in §7.5 and in Chapter 8.

## §6.4 Regions with boundaries

Let us consider briefly what these results imply for a region with a boundary. Obviously if the region is finite then no matter what boundary condition we impose, then the process will become extinct. On an infinite region we can however deduce more.

As an example, suppose that we have a measure-valued diffusion on the upper half plane with normal reflection at the boundary. Consider the following construction. Take a copy of the whole plane with the coefficients of spatial diffusion and branching defined symmetrically about the $x$-axis, and construct the measure-valued diffusion with these coefficients. 'Folding' the lower half plane onto the upper half plane by projecting along the $y$-axis yields the measure-valued diffusion with reflecting boundary conditions. The existence or non-existence of a limiting distribution is then reduced to the corresponding problem for this process on the whole plane. Using §4.5 (or rather its analogue for regions with boundaries) we can obviously extend this 'doubling and projecting' technique considerably.

For a region with absorbing boundary conditions, extinction is clearly 'more likely'. Obviously if the spatial diffusion part of the process is recurrent, then extinction is certain. However, our results do not allow us to deduce anything when the spatial diffusion is transient. The obvious 'test case' is the Dawson process on a half-space with absorbing boundary condition. It is an easy exercise to check that even if the branching coefficient is set to zero, this process will suffer local extinction

in any dimension and obviously then the same is true for any process equivalent to the Dawson process in the sense that we can apply §4.5. No result has yet been established for the general case.

## Chapter 7. Measure-valued Diffusions with Immigration

We saw in Chapter 6 that if we start from a finite initial mass, then our measure-valued diffusion processes will suffer extinction and that this can happen even for an infinite initial mass. However, it would still be of interest to find some sort of 'natural' limiting spatial distribution of the processes. We will see in §7.1 that we cannot find a scaling function $f(t)$, say, for which $f(t)X(t)$ converges to a non-trivial limit, and so we are led to seek some means of 'keeping the processes alive'. Mathematically at least, the obvious thing to try is conditioning on non-extinction. However, this is physically unappealing since we would be conditioning on an event whose probability is tending to zero as time increases and so we turn to the biological models themselves for inspiration. Recall from Chapter 2 that our branching diffusions may model for example the spread of a rare mutant gene in a large normal background or of a small epidemic in a large susceptible population. As we remarked there, in either case, one might expect that rather than a single injection of the rare individuals or infectives into the population, there would be a continuous process of creation (at a possibly time-varying rate). This leads us to consider a discrete particle branching diffusion with a process of 'immigration'. We take this process of immigration to be a Poisson process of mass creation. In §7.2 we find a measure-valued diffusion approximation to these discrete particle processes under a rescaling compatible with that used to construct the measure-valued diffusions without immigration. We show that this diffusion approximation can also be characterized via a deterministic dual process which is obtained simply by integrating the solution to the usual evolution equation with respect to time.

In §7.3, we consider the total mass of a finite measure-valued diffusion process with immigration and write down a rate of immigration for which the total mass of the measure-valued process with immigration converges to a non-trivial limit.

If we are modelling the spread of a rare mutant population, it is most likely that we shall be modelling a constant rate of immigration. Although in this case the process explodes in the limit as time tends to infinity, unlike the process without immigration the process with a constant rate of immigration can be rescaled. The simple calculation in §7.3 shows that if $X^I(t)$ denotes the process with immigration, then $|X^I(t)|/t$ converges in distribution to a non-trivial limit.

For a critical Galton-watson process $\{Z(n)\}_{n\geq 1}$, it is well known that the effect on the limit law of $Z(n)/n$ of immigrating at the constant rate $\sigma^2/2$ (where $\sigma^2$ is the variance of the branching mechanism of the process) is the same as that of conditioning the process without immigration on non-extinction (see, for example, Athreya&Ney (1972) Chapter VI, §7). In §7.4 we prove the corresponding result for our measure-valued diffusions with immigration on finite regions of $\mathbb{R}^d$ and in §7.5 we compare the two effects on processes on infinite regions. We will see that further to the dichotomy already established in Chapter 6 between processes with transient and recurrent spatial diffusions, there is a division of behaviour of the processes with null recurrent and positive recurrent spatial diffusions.

## §7.1 Scaling functions

Even 'though a measure-valued diffusion started from some finite initial measure $\mu$ suffers extinction, we might hope to glean some information from observing 'how the process dies'. For example, we might hope to find a function $f(t)$ for which $f(t)X(t)$ converges to some non-trivial limit as $t\to\infty$. However, it is easy to see that this is

impossible. As usual we set $b\equiv 1$. Recall from §6.1 that if we solve the evolution equation for the process with $u(0)\equiv C$ for some constant $C>0$, then

$$u(t)=\frac{C}{Ct+1}$$

Thus for a finite initial measure $\mu$,

$$\mathbb{E}_\mu[exp(-<1,f(t)X(t)>)]=exp\Big(-\frac{|\mu|f(t)}{tf(t)+1}\Big)$$

so that existence of a function $f$ for which $f(t)X(t)$ has a non-trivial limiting distribution implies existence of a function $f$ for which

$$\frac{f(t)}{tf(t)+1}\to\alpha \qquad \text{for some } 0<\alpha<\infty$$

Therefore we cannot find such a scaling function.

§7.2 The processes with immigration : Definition and characterization

Let us then consider branching diffusion processes with immigration. We begin by describing the discrete particle models. These are branching diffusions with a Poisson process of mass creation. Let us suppose that immigration occurs at a possibly time-variable rate $\rho(t)$ with distribution $\mu$. i.e. The probability that a new particle is created in an element of space $\Delta x$ in the time interval $[t,t+\delta t)$ is $\rho(t)\delta t\mu(\Delta x)+o(\delta t\mu(\Delta x))$. Once inside the domain the particles obey the usual branching diffusion laws. We denote by $\mathbb{E}^{I,1}_{\rho,\mu}$ the expectation operator corresponding to the discrete process with immigration at rate $\rho$ with distribution $\mu$ where the

domain is supposed to be initially empty. We can obviously use the multiplicativity of the process to generalize this to the case where there is already mass in the domain at time zero.

It will be convenient to write

$$G_t(f|x)=\mathbb{E}_{\delta_x}[exp(<logf,X(t)>)] \qquad f\varepsilon\, C_b(\mathbb{R}^d),\ 0<f<1 \tag{7.1}$$

Let us begin with the case when $\rho\equiv$constant and the domain is finite. Then

$$\mathbb{E}^{I,1}_{\rho,\mu}[exp<logf,X(t)>]=\mathbb{E}^{I,1}_{\rho,\mu}\Big(\mathbb{E}^{I,1}_{\rho,\mu}[exp(<logf,X(t)>)|\textit{there are } n \textit{ immigrants in } [0,t) \textit{ with times and locations of immigration } (x_1,s_1),\ \ldots\ (x_n,s_n)]\Big)$$

$$=\mathbb{E}^{I,1}_{\rho,\mu}[\prod_{i=1}^{n} G_{t-s_i}(f|x_i)]$$

$$=\sum_{n=0}^{\infty} \frac{exp(-\rho t\cdot\mu(G))(\rho t\cdot\mu(G))^n}{n!}\cdot\prod_{i=1}^{n}\int_0^t\int_G G_{t-s_i}(f|x_i)\ \frac{d\mu(x_i)}{\mu(G)}\ \frac{ds_i}{t} \tag{7.2}$$

$$=exp(-\rho t\cdot\mu(G))\cdot exp(\rho\int_0^t\int_G G_{t-s}(f|x)\,d\mu(x)\,ds)$$

$$=exp(-\rho\int_0^t\int_G\Big(1-G_s(f|x)\Big)d\mu(x)\,ds) \tag{7.3}$$

If the rate of immigration is not constant, then $\rho t$ in (7.2) is replaced by $\int_0^t\rho(s)ds$ and $\frac{ds_i}{t}$ is replaced by

$$\frac{\rho(s_i)ds_i}{\int_0^t\rho(s)ds}$$

Equation (7.3) then becomes

$$exp(-\int_0^t\int_G \rho(t-s)\Big(1-G_s(f|x)\Big)d\mu(x)\,ds) \tag{7.4}$$

We can also overcome the restriction of a finite domain by considering the effect of immigrating into an increasing sequence of domains $G_n$ and taking the weak limit of the processes obtained in this way as $n\rightarrow\infty$. Equation (7.4) remains unchanged under this limit.

Remark As usual we have tacitly assumed the existence of the processes, but as before we could use the one-dimensional distributions derived above to define the processes and then appeal to the general theory of Markov processes to confirm that they really are the one-dimensional distributions of a strong Markov process.

We now turn to the problem of constructing a diffusion approximation to these discrete particle processes with immigration. We should like to use a rescaling which is compatible with the Feller rescaling used for the processes without immigration. At the $n^{th}$ stage we therefore take the mass of each particle to be $1/n$ and the branching rate to be $nV(x)$, and further we suppose the immigration to be taking place at rate $n\rho$. Let $\mathbb{E}^{I,n}_{\rho,\mu}$ denote the expectation operator of the $n^{th}$ rescaling and let $G^n_s(f|x)$ be the obvious extension of (7.1). Then

$$\mathbb{E}^{I,n}_{\rho,\mu}[exp(<logf_n,X(t)>)]=exp(-\int_0^t\int_G n\rho(t-s)\Big(1-\{G^n_s(f_n|x)\}^{1/n}\Big)d\mu(x)\,ds \tag{7.5}$$

Set $f_n=\left(1-\frac{\phi}{n}\right)^n$, so that

$$exp(<logf_n,X(t)>)\rightarrow exp(-<\phi,X(t)>)$$

Now, from the convergence of the sequence of processes without immigration

$$G_s^n(f_n|x)\rightarrow \mathbb{E}_{\delta_x}[exp(-<\phi,X(s)>)]$$

$$=exp(-u(s,x))$$

where $u(s,x)$ solves the usual evolution equation (4.5). Therefore the right hand side of (7.5) converges and

$$r.h.s.\ of\ (7.5)\rightarrow exp\left(-\int_0^t\int_G \rho(t-s)u(s,x)d\mu ds\right) \qquad as\ n\rightarrow\infty$$

Denoting the expectation operator of the limiting process (if it exists) by $\mathbb{E}^I_{\rho,\mu}$, we see that

$$\mathbb{E}^I_{\rho,\mu}[exp(-<\phi,X(t)>)]=exp\left(-\int_0^t \rho(t-s)<u(s),\mu>ds\right) \tag{7.6}$$

where $u(s)$ solves the evolution equation for the process without immigration. Of course, we should really follow the same strategy as that adopted in Chapter 4 and prove that the sequence of processes obtained under this rescaling is relatively weakly compact. Granted this, we then know that the distribution of any limit point satisfies (7.6) and thus, since we have already established uniqueness of the solution $u(t)$ to

the evolution equation for the process without immigration, such a limit point must be unique. The details of relative compactness are omitted. They are very similar to those given in Chapter 4.

### §7.3 The total mass of the process with immigration

Having constructed the measure-valued process with immigration, let us now establish whether it is worth pursuing this line of investigation in search of a non-trivial limiting distribution. We devote this section to the total mass process of a finite measure-valued branching diffusion process with immigration. As we have already seen, setting $u(0) \equiv C$ in the evolution equation yields

$$u(t) = \frac{C}{Ct+1}$$

Thus the expected total mass at time $t$ of the process with immigration, denoted $|X^I(t)|$, is given by

$$\mathbb{E}[exp(-C|X^I(t)|)] = exp\left(-\int_0^t \left(\frac{C\rho(t-s)}{Cs+1}\right) ds\right) \tag{7.7}$$

If we choose, for example, $\rho(t) = (log(t+2))^{-1}$, then it is not difficult to check (using, for example, the sandwich principle) that this expression converges as $t \to \infty$ to a finite limit (in fact, the limit is 1 if $C=1$ in this case). Thus there do exist rates of immigration for which the total mass of the process converges to a non-trivial, finite limit and we could reasonably expect in such cases that the processes with immigration might have a non-trivial limiting distribution. However, the rates of immigration for which this occurs seem somewhat contrived and at least from a

physical point of view not very natural—if we are modelling the spread of a rare mutant gene we would be much more likely to consider a constant rate of immigration. The rest of the chapter will be devoted to this special case.

From (7.7), it is clear that the mass of a measure-valued diffusion with a constant rate of immigration is unbounded in the limit as $t \to \infty$. However, in this case we can find a rescaling : setting $C=1/t$ in (7.7), we see that the total mass of $X(t)/t$ converges in distribution to a non-trivial finite limit as $t \to \infty$.

## §7.4 Conditioning on non-extinction : Finite regions

Of course, the existence of a non-trivial limiting distribution for the total mass of a rescaled measure-valued diffusion does not necessarily imply the existence of a non-trivial limit for the process itself. In this section we consider such processes on finite regions of $\mathbb{R}^d$ when the distribution of immigration is chosen to be a non-trivial invariant measure of their spatial diffusion part. Using the asymptotic evaluation of the coefficients of the series solution to the evolution equation for the process without immigration which was obtained in Chapter 5 we show that, in this case, not only do we have a non-trivial limiting distribution for these rescaled processes, but also that they provide a nice physical interpretation of conditioning the processes without immigration on non-extinction.

Consider then

$$\mathbb{E}^I_{\rho,\mu}[exp\Big(-\Big\langle \lambda\phi, \frac{X(t)}{t}\Big\rangle\Big)] = exp\Big(-\int_0^t <u_{\lambda/t}(s), \mu> ds\Big)$$

where $u_{\lambda/t}$ solves the usual evolution equation with $u_{\lambda/t}(0) \equiv \lambda/t$. In the notation of Chapter 5, at least if $\lambda/t$ is sufficiently small, we may write this

(7.8) $$\mathbb{E}^I_{\rho,\mu}[exp\Big(-\Big\langle\lambda\phi,\frac{X(t)}{t}\Big\rangle\Big)]=exp\Big(-\int_0^t\rho\Big\langle\sum_{\tau\in\mathcal{T}}\Big(\frac{\lambda}{t}\Big)^{|\tau|}U^\tau_s\phi,\mu\Big\rangle ds\Big)$$

Now, if $\lambda\int\phi d\mu<\frac{1}{8}$, then the series on the right is uniformly convergent in $t$ so that we can interchange the order of integration and summation and the asymptotic expressions of Chapter 5 then give an asymptotic expression for the right hand side of (7.8).

(7.9) $$\mathbb{E}^I_{\rho,\mu}[exp\Big(-\Big\langle\lambda\phi,\frac{X(t)}{t}\Big\rangle\Big)]\sim exp\Big(\rho\sum_{n=1}^{\infty}\frac{1}{n}\Big(-\lambda\int\phi d\mu\Big)^n\Big)$$

We can now use this expression to calculate the moments of the rescaled process with immigration. Our next aim is to show that if $\rho$ is appropriately chosen, then the moments of this process are the same as those of the process without immigration conditioned on non-extinction.

Let us consider then the process without immigration conditioned on non-extinction. We will need to know the value of $\mathbb{P}_\mu[<1,X(t)>\neq 0]$.

$$\mathbb{P}_\mu[<1,X(t)>=0]=\lim_{r\to\infty}\mathbb{E}_\mu[exp(-<r,X(t)>)]$$

Solving the evolution equation with $u(0)\equiv r$ and allowing $r$ to tend to infinity yields

$$\mathbb{P}_\mu[<1,X(t)>\neq 0]=1-\lim_{r\to\infty}exp\Big(-\frac{r\mu(G)}{rt+1}\Big)=\frac{\mu(G)}{t}+o\Big(\frac{1}{t}\Big)$$

Thus for $n\geq 1$,

$$\mathbb{E}_\mu[\left\langle \phi, \frac{X(t)}{t}\right\rangle^n | <1, X(t)>\neq 0] \sim \frac{t}{\mu(G)} \mathbb{E}_\mu[\left\langle \phi, \frac{X(t)}{t}\right\rangle^n]$$

Now, the moments on the right are calculated from

$$\mathbb{E}_\mu[\left\langle \lambda\phi, \frac{\mathrm{X(t)}}{\mathrm{t}}\right\rangle^n] = exp\Big(-\Big\langle \sum_{\tau \varepsilon \mathcal{T}} \Big(\frac{\lambda}{t}\Big)^{|\tau|} \mathrm{U}_t^\tau \phi, \mu\Big\rangle\Big)$$

$$\sim exp\Big(-\frac{\lambda}{t} \cdot \sum_{n=1}^{\infty} \Big(-\lambda \int \phi \, d\mu\Big)^{n-1} \int \phi \, d\mu\Big)$$

Suppose then that we choose $\rho\mu(G)=1$. Differentiating in $\lambda$ and evaluating at $\lambda=0$, it is not difficult now to check that

$$\mathbb{E}_\mu[\left\langle \phi, \frac{\mathrm{X(t)}}{\mathrm{t}}\right\rangle^n | <1, X(t)>\neq 0] \sim \mathbb{E}^I_{\rho,\mu}[\left\langle \phi, \frac{X(t)}{t}\right\rangle^n] \qquad \forall n \geq 0 \tag{7.10}$$

The important thing to notice is that for each fixed $\phi$, these moments satisfy the Carleman condition (see Appendix E) so that they uniquely determine a probability distribution. From this we can then deduce that (7.10) actually implies the following

Theorem 7.1 Suppose that $X(t)$ is a measure-valued critical branching process on a finite region $G \subseteq \mathbb{R}^d$ with reflection on the boundary $\partial G$ of $G$. Then the effect on the limit law of $X(t)/t$ of conditioning on non-extinction is the same as that of immigrating at rate $1/\mu(G)$.

Remark By our definition of $b$, that this is the correct value of $\rho$ could be predicted from the corresponding result for the Galton-Watson process described at the

beginning of the chapter.

We have therefore shown that if we are working on bounded domains in $\mathbb{R}^d$, then the mathematically appealing idea of conditioning the measure-valued processes on non-extinction is, in a sense, equivalent to the physically appealing notion of immigration.

## §7.5 Conditioning on non-extinction : Infinite regions

Obviously Theorem (7.1) only makes sense if $\mu(G)<\infty$. On infinite regions the behaviour can be totally different. We already know that it does not make sense to consider the process with a constant rate of immigration if the spatial diffusion part is transient. What we will observe is a further dichotomy between those processes with null and positive recurrent spatial diffusion parts. In both cases, as we know from Chapter 6, the measure-valued process without immigration suffers local extinction (i.e. the amount of mass in each compact set converges to zero in distribution as time tends to infinity). In this section we investigate the local effects of immigration and of conditioning on non-extinction. As for the case of a finite domain we suppose that mass is immigrating at a constant rate with distribution $\mu$ where $\mu$ is an invariant measure of the spatial diffusion part of the process. We take $\phi=\chi_K$, the indicator function of the compact set $K\subseteq\mathbb{R}^d$ and write as usual $T(t)$ for the maximum expected sojourn time in the set $K$ up to time $t$ of a particle moving under the spatial diffusion part of the branching process (the maximum being taken over all initial positions of the particle).

Let us consider first what happens if the spatial diffusion part of the process is null recurrent. It is easy to see the effect of immigration. Mathematically, what happens in this case is that $T(t)/t\to 0$ as $t\to\infty$, so that (asymptotically) only the first

term of the series expansion of the solution to the evolution equation will contribute to the right hand side of (7.8). Thus, in the limit as $t\to\infty$, all moments other than the first vanish, and

$$\lim_{n\to\infty} \mathbb{E}^I_{\rho,\mu}[\left\langle \phi, \frac{X(t)}{t} \right\rangle] = \rho \int \phi \, d\mu$$

On an infinite region, it is no longer straightforward to condition on non-extinction. What we will do is condition on the process not having suffered local extinction in some compact set $D\subseteq\mathbb{R}^d$ with $K\subseteq D$ by conditioning on the event $[<\psi,X(t)>\neq 0]$ where $\psi=\chi_D$.

Remark In the case of a null recurrent diffusion the behaviour of the resulting process will be independent of $D$, so that we really can regard this as conditioning on 'local non-extinction'.

The difficulty with the process conditioned on this type of non-extinction is that the spatial motion has much more freedom to 'smear the mass out'. Calculation of

$$\mathbb{E}[\left\langle \phi, \frac{X(t)}{t} \right\rangle | <\psi, X(t)> \neq 0]$$

will be straightforward once we have calculated $\mathbb{P}[<\psi,X(t)>\neq 0]$. Recall from Chapter 5 that

$$\beta \tilde{T}(t)\left(\int \psi \, d\mu\right)^2 \leq \mathbb{E}[<\phi,X(t)>^2] \leq \tilde{T}(t)\left(\int \psi \, d\mu\right)^2$$

for some $\beta>0$ where $\tilde{T}(t)$ is the maximum expected sojourn time in the set $D$ of a particle moving under the spatial diffusion part of the process. Now, it is clear that

$$\mathbb{P}[<\psi,X(t)>>\tilde{T}(t)]\leq\frac{\mathbb{E}[<\psi,X(t)>]}{\tilde{T}(t)}=\frac{\mu(D)}{\tilde{T}(t)} \tag{7.11}$$

Also, from Chapter 6

$$\mathbb{E}[exp(-<\psi,X(t)>)]\rightarrow 1 \qquad \text{as } t\rightarrow\infty$$

and conditioning this on the event $[<\psi,X(t)><\tilde{T}(t)]$ and using (7.11) we see that

$$\mathbb{E}[exp(-<\psi,X(t)>)|<\psi,X(t)><\tilde{T}(t)]\rightarrow 1 \qquad \text{as } t\rightarrow\infty \tag{7.12}$$

Now since $\tilde{T}(t)$ is finite for each fixed $t$, we can express the exponential as a sum and interchange the order of integration and summation in (7.12) which leads to

$$\mathbb{E}[<\psi,X(t)>^n|<\psi,X(t)><\tilde{T}(t)]\rightarrow 0 \qquad \text{as } t\rightarrow\infty,\ \forall n\geq 1$$

Thus

$$\mathbb{P}[<\psi,X(t)>\neq 0]\sim\mathbb{P}[<\psi,X(t)>>\tilde{T}(t)]$$

and for each $n\geq 1$,

$$\mathbb{E}[<\psi,X(t)>^n|<\psi,X(t)>\neq 0]\sim\mathbb{E}[<\psi,X(t)>^n|<\psi,X(t)>>\tilde{T}(t)] \tag{7.13}$$

Using the expression for the first order moment of the process obtained in Chapter 6 and (7.13) to find an (asymptotic) expression for the variance of the process conditioned on the event $[<\phi,X(t)>>\tilde{T}(t)]$ which is necessarily a non-negative quantity yields finally

$$\mathbb{P}[<\psi,X(t)>\neq 0]\geq\frac{\mu(D)}{\beta\tilde{T}(t)} \qquad \text{as } t\rightarrow\infty$$

and then

$$\mathbb{E}[\left\langle\phi,\frac{X(t)}{t}\right\rangle|<\psi,X(t)>\neq 0]\leq\frac{\mu(K)\tilde{T}(t)\beta}{\mu(D)t} \qquad \text{as } t\rightarrow\infty$$

No matter how large the set $D$, we see that this expression tends to zero. This is a reflection of the tendency of mass to cluster outside any bounded set.

The case of a positive recurrent diffusion is a straightforward extension of §7.4 and so we omit the proof. We are now once more in a position to condition on global non-extinction (since $|\mu|$ is finite), and since $\mu$ is tight this actually is the same as conditioning on $[<\psi,X(t)>\neq 0]$ and allowing $D$ to increase to $\mathbb{R}^d$; i.e. local and global non-extinction are equivalent. Mathematically, what is different in this case from the null recurrent case is that $T(t)/t\not\rightarrow 0$ as $t\rightarrow\infty$. Physically, the diffusion no longer serves to push the mass out of finite sets.

Thus on infinite regions, for the process with immigration to have the nice physical interpretation of Theorem (7.1), it must have a positive recurrent spatial diffusion part.

## Chapter 8. Related Results and Open Problems

The study of measure-valued diffusion processes is still in its infancy and it seems wrong to close any work on the subject with a chapter entitled 'conclusion'. The rôle of this closing chapter is rather to indicate some results related to those presented here and indicate a number of open questions. It is worth mentioning that all the measure-valued processes which we have studied belong to the class of processes which Dynkin (1988) calls *superprocesses*. We have avoided this term since it is unclear whether or not it will survive. [My personal view is that we must have some term for this class of processes and while I would prefer a process $X$ to be called a superprocess over a process $Y$ only if $Y$ were a subprocess of $X$, I have yet to see any better suggestions!]

We begin by fulfilling our promises and discussing briefly the effects of dropping the assumptions of §3.1 and §5.2. Recall that in §3.1, we made a number of regularity assumptions. Essentially these were to ensure that the spatial diffusion part of the process should be 'well-behaved' and that the branching part should be non-vanishing, continuous and bounded. For the purposes of constructing the processes these could be relaxed substantially. For example, we could certainly allow the branching part to vanish, and we only require the spatial motion to be sufficiently nice that we can prove a 'compact containment' condition for the rescaled processes (c.f. Chapter 4). The difficulties arise in investigating asymptotic behaviour. We cannot really expect the measure-valued processes to be well-behaved asymptotically if their spatial diffusion part alone is badly behaved! Although we could relax the condition on the form of the generator, we must insist on Assumption (v). In our investigation of the

asymptotic behaviour we also made the assumption that the branching mechanism should be constant. For a recurrent spatial diffusion we can obviously extend this to any branching mechanism satisfying the conditions of §3.1, for we simply bound the solution to the evolution equation for the process by that for the process whose branching part is everywhere given by the infimum of the branching mechanism. Extinction follows. Since the branching mechanism will only vanish at a point if either a particle of our discrete model residing at that point has infinite lifetime or if at that point the offspring generating function of each particle has zero variance which implies that each particle has one offspring with probability one, this problem is of very little biological significance. However, if, as mathematicians, we do not allow this to deter us, our original proof suggests an obvious means of attack—we should investigate the effect on the second term of our series solution to the evolution equation of choosing such a branching mechanism. (If the branching mechanism is strictly positive on a set of infinite $\mu$-measure then it is easy to see that the process will still suffer extinction—thus in particular we could drop the assumption that there was a uniform lower bound on the branching mechanism and simply assert that it should be positive. If the branching mechanism is only positive on a bounded set then the question remains unresolved.) In the transient case also we could drop the assumption that the branching mechansim is constant, and in fact in this case we don't even have to worry about the branching part vanishing.

An important assumption which we should discuss is that of criticality of the branching mechanism. As we remarked at the beginning of Chapter 4, this assumption is essential to the convergence of the processes under our rescaling. However, it is possible to construct 'supercritical' measure-valued branching processes in a very similar way. Specifically, we take a sequence $\{Z^n\}_{n\geq 1}$ of discrete particle

branching diffusion processes all with the same spatial diffusion part (with generator $L$, satisfying the usual regularity conditions) and with continuous branching rates and mechanisms given by $V_n(x)$, $\Phi_n(s,x)$ respectively. Then provided that

$$a(x)=\lim_{n\to\infty} V_n(x)\Big(\Phi_n(1,x)-1\Big)$$

$$b(x)=\lim_{n\to\infty}\frac{1}{2n}V_n(x)\frac{\partial^2}{\partial s^2}\Phi(s,x)\Big|_{s=1}$$

are bounded functions of $x$, the sequence $\{\frac{1}{n}Z^n\}_{n\geq 1}$ converges weakly to the measure-valued limit $X(t)$, again characterized via an evolution equation

$$\mathbb{E}_\mu[exp(-<\phi,X(t)>)]=exp(-<u(t),\mu>)$$

where

$$\frac{\partial u}{\partial t}=Lu+au-bu^2$$

$$u(0)=\phi$$

The total mass of this process on $\mathbb{R}^d$ (or on a region of $\mathbb{R}^d$ with reflecting boundary conditions) is a one-dimensional diffusion with generator

$$A=bx\frac{d^2}{dx^2}+ax\frac{d}{dx}$$

which has extinction probability

$$exp\Big(-\frac{aX(0)}{b}\Big)$$

(see Bailey (1964) Chapter 14). Even on a bounded region with absorbing boundary the extinction probability may be less than one (see Wang (1982)). For this type of process we can therefore expect interesting limiting behaviour (at least for sufficiently large $a$). Unfortunately, we would not expect such a limiting distribution to throw any light on 'natural limits' of our critical processes, since the effect of the supercriticality term $a$ is, of course, that of a process of mass creation where the amount of mass created in a given set in the time interval $[t,t+\delta t)$ is proportional to the amount of mass already there—thus amplifying the discrepancies between quantities of mass in disjoint sets. These processes are of course worthy of study in their own right, and at least in principal we could write down the solution to their evolution equation as a Wild sum, 'though calculation of the coefficients would obviously be more difficult than in the critical case.

The reader should by this stage have realized that the representation of the deterministic dual process as the solution to the evolution equation is deceptively simple. Indeed, even if we could find a nice closed form for this solution, its application is limited for we only immediately yield results about convergence in distribution of the measure-valued processes. However, with more work we can deduce a little more. Using this dual, Iscoe (1986b) characterized the distribution of the 'weighted occupation time', $Y(t)$, for the Dawson process.

$$<\phi, Y(t)> = \int_0^t <\phi, X(s)> ds \qquad \phi \varepsilon \mathcal{D}(L)$$

The characterization is again via a deterministic dual process

$$\mathbb{E}_\mu[exp(-<\phi, Y(t)>)] = exp(-<v(t), \mu>)$$

where

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v - v^2 + \phi$$

$$v(0) = 0$$

(In fact he proves this result in rather greater generality—the spatial diffusion part of the process can be any symmetric stable process and the branching mechanism can have infinite variance.) It is not difficult to check that this result generalizes to any of our measure-valued processes. In Iscoe (1986b), it is established that for the Dawson process on $\mathbb{R}^1$ with initial value Lebesgue measure, the weighted occupation time in a compact set of the process remains finite. In Iscoe (1986a) this result is strengthened to prove that we have a.s. convergence to zero of the amount of mass in a compact set. I believe that it has been established by Stanley Sawyer (personal communication via Ted Cox) that for the two-dimensional critical branching random walk (the lattice analogue of the two-dimensional Dawson process), the weighted occupation time in any compact set tends to infinity as time increases. This suggests that the same is true of the Dawson process in two dimensions, so that one would not expect a.s. convergence in this case. It would be interesting to establish a criterion for a.s. convergence for those processes with a general recurrent spatial diffusion part. It seems likely that there will be some critical value of $\rho$ in Assumption (v) of §3.1 above which we will have a.s. convergence.

There is a general philosophy with this type of measure-valued process that a result is true for the measure-valued process if and only if it is true for the discrete process to which it is a diffusion approximation. (From the biological point of view, this is of course the real value of studying the measure-valued processes!) Indeed from their very construction the measure-valued processes lend themselves to the techniques of non-standard analysis and this has been exploited by Perkins (1988) in investigating the way in which the Dawson process spreads its mass over its support. This support was first studied by Dawson&Hochberg (1979), where using the rescaling properties of the approximating sequence of branching diffusions it was established that in dimensions two and above, at any given positive time, the Dawson process started from Lebesgue measure is a singular measure. Dawson, Iscoe&Perkins (1988) have even established the exact Hausdorff measure function of the support process, but this still does not provide us with a concrete description of the support of such a process. If there were such a description then it must surely lead to a much deeper intuition for these processes.

We mention briefly some of the many questions arising from the loose ends in Chapter 7. An obvious open problem is to establish the nature of the limit of the process if we choose the rate of immigration in such a way that the total mass converges. Another intriguing question is the following. We know that the Dawson process on a half space suffers extinction, but are there spatial diffusions for which the measure-valued processes will not suffer extinction? For example, what happens if there is a drift away from the boundary of the space?—this would not only increase the amount of mass 'at infinity' initially, but would also decrease the probability of mass hitting the boundary and consequently disappearing. Perhaps a more interesting problem is to establish how much immigration it would require to keep a

process with absorbing boundary conditions alive.

There is certainly a wealth of open problems related to the processes which we have studied in this thesis, and although, as we have seen, we can deduce a lot from their characterization via our deterministic dual it is clear that we would benefit greatly from the discovery of a more manageable characterization or the development of more satisfactory ways of dealing with the evolution equation. It is certainly the case, for example, that the processes have other duals (in the sense of Theorem C.1). The other obvious approach is to use computer simulations, but these are necessarily discrete models and start from a finite measure. Although this second problem could be overcome to an extent by generalizing the work of Cox (1988) which establishes for a similar model how long the process must run before 'it realises that it is finite', we should not lose sight of the fact that these processes will only be of lasting value to the biologists if they enable us to deduce results which have not already been proved for the discrete processes.

## Appendix A. The weak topology on $\mathcal{M}_0(\mathbb{R}^d)$

Recall that a sequence of measures $\{\mu_n\}_{n\geq 1}$ on $\mathbb{R}^d$ is said to converge weakly to the limit $\mu$ iff

$$\lim_{n\to\infty}\int f d\mu_n = \int f d\mu \qquad \forall f \in C_b(\mathbb{R}^d)$$

We describe here a metric for the weak topology on $\mathcal{M}_0(\mathbb{R}^d)$ (the space of finite, positive Borel measures on $\mathbb{R}^d$). For this we will require the Prohorov metric on $\mathcal{P}(\mathbb{R}^d)$—the family of Borel probability measures on $\mathbb{R}^d$. Let $\mathcal{C}$ denote the collection of closed subsets of $\mathbb{R}^d$, and for $F\in\mathcal{C}$ and $\epsilon>0$ define

$$F^\epsilon = \{x\in\mathbb{R}^d: \inf_{y\in F} d(x,y)<\epsilon\}$$

($d$ here is the ordinary Euclidean metric on $\mathbb{R}^d$).

Definition A.1 The *Prohorov metric*, $\tilde{\rho}$, on probability measures is defined as follows. For $P$, $Q\in\mathcal{P}(\mathbb{R}^d)$

$$\tilde{\rho}(P,Q) = inf\{\epsilon>0: P(F)\leq Q(F^\epsilon)+\epsilon \ \forall F\in\mathcal{C}\} \tag{A.1}$$

The reader is referred to Ethier&Kurtz (1986), Chapter 3, §1 for a proof that this does indeed define a metric on $\mathcal{P}(\mathbb{R}^d)$.

This suggests an obvious way to define a metric on the non-zero elements of $\mathcal{M}_0(\mathbb{R}^d)$. As proposed in Ethier&Kurtz (1986) Chapter 9 Exercise 7, we define

$$\rho(\mu,\nu)=\tilde{\rho}\left(\frac{\mu}{|\mu|},\frac{\nu}{|\nu|}\right)+\big||\mu|-|\nu|\big| \tag{A.2}$$

where $|\mu|=\int 1\,d\mu$. That this defines a metric on the non-zero measures of $\mathcal{M}_0(\mathbb{R}^d)$ is a trivial verification. The completion of this metric space, denoted $\overline{\mathcal{M}_0(\mathbb{R}^d)}^{\rho}$, can be identified with $\mathcal{M}_0(\mathbb{R}^d)$. Obviously this metric makes $\mathcal{M}_0(\mathbb{R}^d)$ into a complete separable metric space. We should like to see that the metric $\rho$ induces the weak topology, but this follows easily from the fact that $\tilde{\rho}$ induces the weak topology on $\mathcal{P}(\mathbb{R}^d)$ (see Ethier&Kurtz (1986), Chapter 3 §3).

Remark In fact the topology induced by $\rho$ on $\mathcal{M}_0(\mathbb{R}^d)$ is stronger than the weak topology whereas the topology induced by $\tilde{\rho}$ on $\mathcal{P}(\mathbb{R}^d)$ is equivalent to the weak topology. The difference arises because in taking the completion, we are taking equivalence classes of Cauchy sequences in $\mathcal{M}_0(\mathbb{R}^d)\backslash\{0\}$, where two sequences $\{\mu_n\}_{n\geq 1}$, $\{\nu_n\}_{n\geq 1}$ are deemed to be equivalent if $\rho(\mu_n,\nu_n)\to 0$ as $n\to\infty$. This gives rise then to an infinite number of measures of total mass zero in the completion. We identify all such measures in the weak topology.

## Appendix B. An existence theorem

We state here some results from the general theory of Markov processes sufficient to ensure the existence of our branching diffusion processes of Chapter 3.

Let $(E,\mathcal{B})$ denote an arbitrary topological state space. ($\mathcal{B}$ here is the collection of Borel subsets of $E$.) We write $C_b(E)$ for the set of bounded measurable functions on $E$.

Theorem (see Dynkin (1965) Vol 1, Chapter II, §1)
A necessary and sufficient condition for a contraction semigroup of linear operators $T_t$ on the space $C_b(E)$ to correspond to some transition function $P(t,x,\Gamma)$ on the space $(E,\mathcal{B})$ is that the following conditions be satisfied

(a) If $f(x)\geq 0 \ \forall x\varepsilon E$, then $T_t f(x)\geq 0 \ \forall x\varepsilon E$

(b) If $f(x_0)=0$, $T_0 f(x_0)=0$

A transition function is said to be normal if $P(0_+,x,E)=1 \ \forall x\varepsilon E$. For a point $x\varepsilon E$, write $U_\epsilon(x)^c$ for the complement of the $\epsilon$-neighbourhood of $x$. A semicompact space is a locally compact Hausdorff space with countable base.

Theorem (see Dynkin (1965) Vol 1, Chapter III, §2)

Let $P(t,x,\Gamma)$ be a normal transition function on a semicompact $E$. Suppose that for every compact $\Gamma$

(i) $$\lim_{t\downarrow 0} \sup_{x\in\Gamma} P(t,x,U_\epsilon(x)^c)=0$$

(ii) For each $u\geq 0$,

$$\lim_{y\to\infty} \sup_{t\leq n} P(t,y,\Gamma)=0$$

Then there is a Markov process with transition function $P(t,x,\Gamma)$ which is bounded, right-continuous and has no discontinuities other than jumps.

A transition function $P(t,x,\Gamma)$ is a Feller function if for any bounded continuous function $f$ on $E$,

$$T_t f(x)=\int_E P(t,x,dy)f(y)$$

is continuous. A process is a Feller process if its transition function is a Feller function.

Theorem (see Dynkin Vol 1, Chapter III, §3)

Every right-continuous Feller process on the topological state space $E$ is a strong Markov process.

## Appendix C. The method of duality

We present here the derivation of the simple form of the so-called 'method of duality' which was used in Chapter 4.

Theorem C.1 (The method of duality)

Let $E_1$, $E_2$ be metric spaces and suppose that $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$ are probability distributions on the space of càdlàg functions from $[0,\infty)$ to $E_1$, $E_2$ respectively. (These spaces are endowed with the natural $\sigma$-fields c.f. p.22.) Take two functions $f$, $g$ on $E_1 \times E_2$ for which the following are true:

(i) $$\{\big(f(\cdot,y),g(\cdot,y)\big) : y \varepsilon E_2\} \subseteq C(E_1)\times C(E_1)$$

(ii) $$\{\big(f(x,\cdot),g(x,\cdot)\big) : x \varepsilon E_1\} \subseteq C(E_2)\times C(E_2)$$

(iii) For $y \varepsilon E_2$

$$f(X(t),y) - \int_0^t g(X(s),y)\,ds \text{ is a } \hat{\mathbb{P}}^{X(0)}\text{-martingale}$$

(iv) For $x \varepsilon E_1$

$$f(x,Y(t)) - \int_0^t g(x,Y(s))\,ds \text{ is a } \tilde{\mathbb{P}}^{Y(0)}\text{-martingale}$$

(v) $f$, $g$ are bounded.

Then

(C.1) $$\hat{\mathbb{E}}^{X(0)}[f\big(X(t),Y(0)\big)] = \tilde{\mathbb{E}}^{Y(0)}[f\big(X(0),Y(t)\big)]$$

Proof From these assumptions, if $\mathbb{E}$ denotes the expectation operator corresponding to the direct product of the distributions $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$ we see that

$$\frac{d}{ds}\mathbb{E}[f\Big(X(s),Y(t-s)\Big)]=\mathbb{E}[g\Big(X(s),Y(t-s)\Big)-g\Big(X(s),Y(t-s)\Big)]=0$$

Thus integrating in $s$ over [0,t] we obtain

$$\mathbb{E}[f\Big(X(0),Y(t)\Big)]=\mathbb{E}[f\Big(X(t),Y(0)\Big)]$$

The conclusion of the Theorem is now immediate.

## Appendix D. Ordered binary rooted trees

Although the definitions and results relating to binary rooted trees which we have used in this work are straightforward, they play an important rôle. Therefore, in order to keep this thesis as self-contained as possible, we gather them together here.

We may construct the set of ordered binary rooted trees, $\mathcal{T}$, inductively. We let $\tau_e$ denote the tree consisting of a single vertex. Given two trees $\tau_1$ and $\tau_2$, $\tau_1 \circ \tau_2$ is the tree obtained by connecting the roots of $\tau_1$ and $\tau_2$ to a single vertex. For our purposes we must consider ordered trees i.e. the two trees $\tau_1 \circ \tau_2$ and $\tau_2 \circ \tau_1$ are deemed to be different if $\tau_1 \neq \tau_2$ (c.f. Chapter 5). The order $|\tau|$ of $\tau \varepsilon \mathcal{T}$ is defined to be the number of vertices having at most one neighbour. Thus $|\tau_1 \circ \tau_2| = |\tau_1| + |\tau_2|$. The result which we require is the following.

Theorem (D.1) Let $C(n)$ denote the number of ordered binary rooted trees of order n, then the power series

$$\sum_{n=1}^{\infty} C(n)\theta^n \tag{D.1}$$

has radius of convergence $\frac{1}{4}$.

Proof From the construction described above, given any such tree of order $n \geq 2$ we can decompose it into two trees of lower order (by removing the root). It is easy to see that we have the following recurrence relation

(D.2) $$C(n)=\sum_{j=1}^{n-1} C(j)\,C(k-j)$$

Consider the complex function $f(z)=\sum_{n=1}^{\infty} C(n)z^n$. Suppose that this series converges.

$$(f(z))^2 = \Big(\sum_{n=1}^{\infty} C(n)z^n\Big)^2$$

$$= \sum_{k=1}^{\infty}\sum_{i=1}^{k} C(i)\,C(k+1-i)z^{k+1}$$

$$= \sum_{n=2}^{\infty}\sum_{i=1}^{n} C(i)\,C(n-i)z^n$$

$$= \sum_{n=2}^{\infty} C(n)z^n$$

$$= f(z)-z$$

Thus if the series converges we have

$$(f(z))^2 + z = f(z)$$

Solving this quadratic and choosing f(0)=0 we obtain

$$f(z)=\frac{1}{2}\Big(1-\sqrt{1-4z}\Big)$$

$f(z)$ is holomorphic on the disc of radius $\frac{1}{4}$ and its power series is (D.1) which implies the result.

## Appendix E. The Carleman condition

Let $\{\alpha_0=1,\alpha_1,\alpha_2\ldots\}$ be a sequence of real numbers. Suppose that there is at least one distribution function $F(x)$, $x\varepsilon I$ such that

$$\alpha_n=\int x^n dF(x) \qquad n=0,1,2,\ldots$$

A sufficient (but not necessary) condition for $F$ to be uniquely determined by $\{\alpha_n\}$ is that the following condition be fulfilled

$$\sum_{n=1}^{\infty}\alpha_{2n}^{-1/(2n)}=\infty$$

This condition is known as the Carleman condition, and we say that the moment problem is determined. (See, for example, Stoyanov (1988) §11 for a discussion of the moment problem.)

## References


Athreya K.B.&P.E. Ney (1972). *Branching Processes.* Springer-Verlag.

Bailey N.T.J. (1964). *The Elements of Stochastic Processes with Applications to the Natural Sciences.* Wiley.

Cox J.T. (1988). Coalescing random walks and voter model concensus times on the torus in $Z^d$. *Preprint.*

Csink L.&B.Øksendal (1983). Stochastic harmonic morphisms : functions mapping the paths of one diffusion into the paths of another. *Ann. Inst. Fourier Grenoble* 33 *No.*2 *p.*219-240.

Dawkins R. (1976). *The Selfish Gene.* Oxford University Press.

Dawson D.A. (1975). Stochastic evolution equations and related measure processes. *J. Multivariate Anal.* 5 *p.*1-52.

Dawson D.A. (1977). The critical measure diffusion process. *Z. Warsch. Verw. Gebiete* 40 *p.*125-145.

Dawson D.A.&K.J. Hochberg (1979). The carrying dimension of a measure-diffusion process. *Ann. Prob.* 7 *p.*693-703.

Dawson D.A., I.Iscoe&E.A.Perkins (1988). Super-Brownian motion : Path properties and hitting probabilities. *Preprint.*

Dawson D.A.&T.G.Kurtz (1982). Applications of duality to measure-valued processes. *Lect. Notes Cont. Inf. Sci.* 42 *p.*177-191. Springer-Verlag.

Dynkin E.B. (1965). *Markov Processes. Vol. I.* Springer-Verlag.

Ethier S.N.&T.G. Kurtz (1986). *Markov Processes: Characterization and convergence.* Wiley.

Feller W. (1951). Diffusion processes in genetics. *Proc. 2nd Berkeley Symp. Math. Statist. Prob.* p.227-246. University of California Press.

Feller W. (1966). *An Introduction to Probability Theory and its Applications Vol.II.* Wiley.

Frauenthal J.C. (1980). *Mathematical Modeling in Epidemiology.* Springer-Verlag.

Ikeda N., M.Nagasawa&S.Watanabe (1968). Branching Markov Processes II. *J. Math. Kyoto Univ. No.*8 p.365-410.

Iscoe I. (1986a). On the supports of measure-valued critical branching Brownian motion. *Preprint.*

Iscoe I. (1986b). A weighted occupation time for a class of measure-valued branching processes. *Prob. Th. Rel. Fields* 71 p.85-116.

Kallenberg O. (1975). Stability of critical cluster fields. *Math. Nachr.* 77 p.7-43.

Perkins E.A. (1988). A space-time property of a class of measure-valued branching diffusions. *Trans. A.M.S.* 305 *No.*2 p.743-795.

Rosen G. (1975). Solutions to the non-linear recombination equation for the infinite spatial domain. *Siam. J. Appl. Math.* 29 *No.*1 p.146-151.

Sawyer S.&J. Fleischmann (1978). The maximum geographical range of a mutant allele considered as a subtype of a Brownian branching random field. *Proc. Nat. Acad. Sci. U.S.A.* 76 *No.*2 p.872-875.

Skorohod V.V. (1964). Branching diffusion processes. Translated by B. Seckler. *Th. Prob. Appl. IX No.*3 p.445-449.

Stoyanov J. (1987). *Counterexamples in Probability.* Wiley.

Sugitani S. (1987a). Some properties for the occupation times arising from the measure-valued branching diffusion processes. *Preprint.*

Sugitani S. (1987b). Some properties of the measure-valued branching diffusion processes. *Preprint.*

Watanabe S. (1968). A limit theorem of branching processes and continuous state branching processes. *J.Math. Kyoto Univ.* 8 *No.*1 *p.*141-167.

Wang F.J. (1982). Probabilities of extinction of multiplicative measure-valued diffusions with absorbing boundary. *Indiana Univ. Math. J.* 31*p.*97-107.

Wild E. (1951). On Boltzmann's equation in the kinetic theory of gases. *Proc. Camb. Phil. Soc.* (1951) *p.*602-609.

Dynkin E.B. (1988). Superprocesses and their linear additive functionals. *Preprint.*